

Cloning Scheme for Generating pCMV-NS35

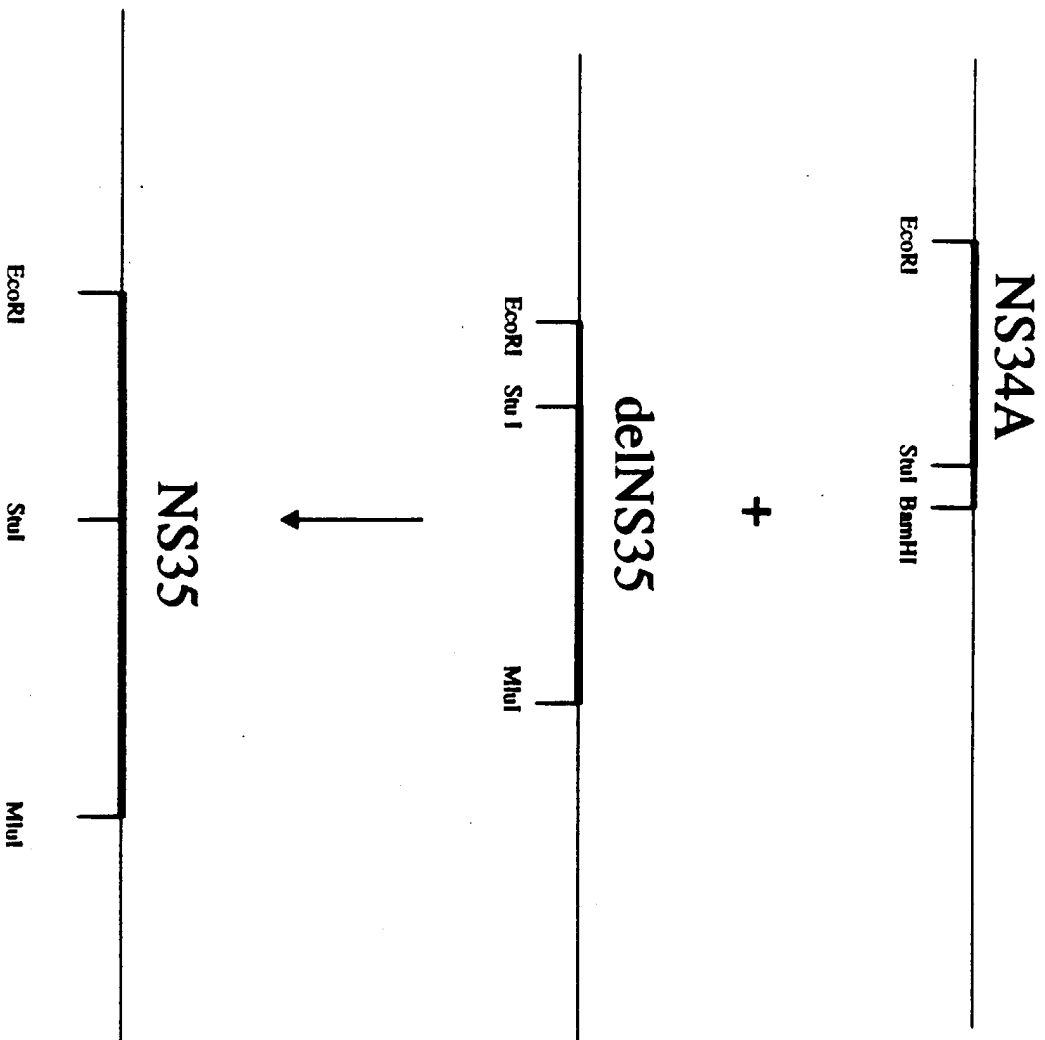
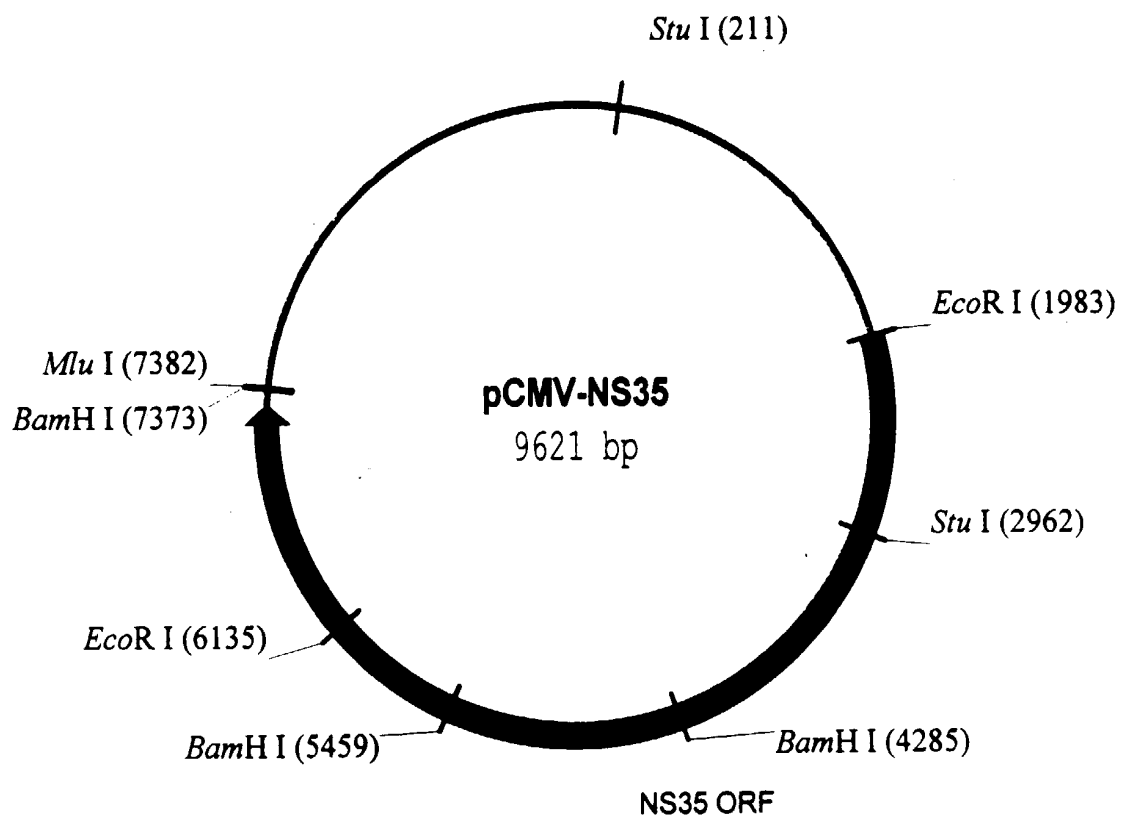


FIGURE 1

FIGURE 2



0022149-41300

FIGURE 3 - Page 1

1	TCGCGCGTTT	CGGTGATGAC	GGTGAAAACC	TCTGACACAT	GCAGCTCCCG	GAGACGGTCA	CAGCTTGTCT	GTAAGCGGAT
	AGCGCGCAAA	GCCACTACTG	CCACTTTTGG	AGACTGTGTA	CGTCGAGGGC	CTCTGCCAGT	GTCGAACAGA	CATTGCGCTA
81	GCCGGGAGCA	GACAAGCCCC	TCAGGGCGCG	TCAGCGGGTG	TTGGCGGGTG	TCGGGGCTGG	CTTAACATATG	CGGCATCAGA
	CGGCCCTCGT	CTGTTCGGGC	AGTCCCGCGC	AGTCGCCCAC	AACCGCCCAC	AGCCCCGACC	GAATTGATAC	GCCGTAGTCT
					StuI	~~~~~		
161	GCAGATTGTA	CTGAGAGTGC	ACCATATGAA	GCTTTTTGCA	AAAGCCTAGG	CCTCCAAAAA	AGCCTCCTCA	CTACTTCTGG
	CGTCTAACAT	GACTCTCACG	TGGTATACTT	CGAAAAACGT	TTTCGGATCC	GGAGGTTTTT	TCGGAGGAGT	GATGAAGACC
241	AATAGCTCAG	AGGCCGAGGC	GGCCTCGGCC	TCTGCATAAA	TAAAAAAAT	TAGTCAGCCA	TGGGGCGGAG	AATGGGCGGA
	TTATCGAGTC	TCCGGGTCCG	CCGGAGCCGG	AGACGTATTT	ATTTTTTTTA	ATCAGTCGGT	ACCCCGCCTC	TTACCCGCTT
321	ACTGGGCGGG	GAGGGAATTA	TTGGCTATTG	GCCATTGCAT	ACGTTGTATC	TATATCATAA	TATGTACATT	TATATTGGCT
	TGACCCGCCC	CTCCCTTAAT	AACCGATAAC	CGGTAACGTA	TGCAACATAG	ATATAGTATT	ATACATGTAA	ATATAACCGA
401	CATGTCCAAT	ATGACCGCCA	TGTTGACATT	GATTATTGAC	TAGTTATTAA	TAGTAATCAA	TTACGGGGTC	ATTAGTTCAT
	GTACAGGTAA	TACTGGCGGT	ACAACTGTAA	CTAATAACTG	ATCAATAATT	ATCATTAGTT	AATGCCCCAG	TAATCAAGTA
481	AGCCCATATA	TGGAGTTCCG	CGTTACATAA	CTTACGGTAA	ATGGCCCGCC	TGGCTGACCG	CCCAACGACC	CCCGCCCAT
	TCGGGTATAT	ACCTCAAGGC	GCAATGTATT	GAATGCCATT	TACCGGGCGG	ACCGACTGGC	GGGTGCTGG	GGGCGGGTAA
561	GACGTCAATA	ATGACGTATG	TTCCCATAGT	AACGCCAATA	GGGACTTTCC	ATTGACGTCA	ATGGGTGGAG	TATTTACGGT
	CTGCAGTTAT	TACTGCATAC	AAGGGTATCA	TTGCGGTTAT	CCCTGAAAGG	TAAGTGCAGT	TACCCACCTC	ATAAATGCCA
641	AAACTGCCCC	CTTGGCAGTA	CATCAAGTGT	ATCATATGCC	AAGTCCGCCC	CCTATTGACG	TCAATGACGG	TAAATGGCCC
	TTTGACGGGT	GAACCGTCAT	GTAGTTCACA	TAGTATACGG	TTCAGGCGGG	GGATAACTGC	AGTTACTGCC	ATTTACCGGG
721	GCCTGGCATT	ATGCCCAGTA	CATGACCTTA	CGGGAAGTTC	CTACTTGGCA	GTACATCTAC	GTATTAGTCA	TCGCTATTAC
	CGGACCGTAA	TACGGGTCAT	GTACTGGAAT	GCCCTGAAAG	GATGAACCGT	CATGTAGATG	CATAATCAGT	AGCGATAATG
801	CATGGTGATG	CGGTTTTTGG	AGTACACCAA	TGGGCGTGGA	TAGCGGTTTG	ACTCACGGGG	ATTTCCAAGT	CTCCACCCCA
	GTACCACTAC	GCCAAAACCG	TCATGTGGTT	ACCGGCACCT	ATCGCCAAAC	TGAGTGCCCC	TAAAGGTTCA	GAGGTGGGGT
881	TTGACGTCAA	TGGGAGTTTG	TTTTGGCACC	AAAATCAACG	GGACTTTCCA	AAATGTCGTA	ATAACCCCGC	CCCGTTGACG
	AACTGCAGTT	ACCCTCAAAC	AAAACCGTGG	TTTTAGTTGC	CCTGAAAGGT	TTTACAGCAT	TATTGGGGCG	GGGCAACTGC
961	CAAATGGGCG	GTAGGCGTGT	ACGGTGGGAG	GTCTATATAA	GCAGAGCTCG	TTTAGTGAAC	CGTCAGATCG	CCTGGAGACG
	GTTTACCCGC	CATCCGCACA	TGCCACCCTC	CAGATATATT	CGTCTCGAGC	AAATCACTTG	GCAGTCTAGC	GGACCTCTGC
1041	CCATCCACGC	TGTTTTGACC	TCCATAGAAG	ACACCGGGAC	CGATCCAGCC	TCCGCGCCCG	GGAACGGTGC	ATTGGAACGC
	GGTAGGTGCG	ACAAAACCTG	AGGTATCTTC	TGTGGCCCTG	GCTAGGTCCG	AGGCGCCGGC	CCTTGCCACG	TAACCTTGCG
1121	GGATTCCCCG	TGCCAAGAGT	GACGTAAGTA	CCGCCTATAG	ACTCTATAGG	CACACCCCTT	TGGCTCTTAT	GCATGCTATA
	CCTAAGGGGC	ACGGTTCTCA	CTGCATTCTT	GGCGGATATC	TGAGATATCC	GTGTGGGGAA	ACCGAGAATA	CGTACGATAT
1201	CTGTTTTTGG	CTTGGGGCCT	ATACACCCCC	GCTCCTTATG	CTATAGGTGA	TGGTATAGCT	TAGCCTATAG	GTGTGGGTAA
	GACAAAAACC	GAACCCCGGA	TATGTGGGGG	CGAGGAATAC	GATATCCACT	ACCATATCGA	ATCGGATATC	CACACCCAAT
1281	TTGACCATTA	TTGACCACTC	CCCTATTGGT	GACGATACTT	TCCATTACTA	ATCCATAACA	TGGCTCTTTG	CCACAACAT
	AACTGGTAAT	AACTGGTGAG	GGGATAACCA	CTGCTATGAA	AGGTAATGAT	TAGGTATTGT	ACCGAGAAAC	GGTGTGATA
1361	CTCTATTGGC	TATATGCCAA	TACTCTGTCC	TTCAGAGACT	GACACGGACT	CTGTATTTTT	ACAGGATGGG	GTCCATTTAT
	GAGATAACCG	ATATACGGTT	ATGAGACAGG	AAGTCTCTGA	CTGTGCCTGA	GACATAAAAA	TGTCCTACCC	CAGGTAAATA

0972449.1220

0972449.1220

FIGURE 3 - Page 2

1441 TATTTACAAA TTCACATATA CAACAACGCC GTCCCCCGTG CCCGCAGTTT TTATTAAACA TAGCGTGGGA TCTCCGACAT
ATAAATGTTT AAGTGATAT GTTGTGCGG CAGGGGGCAC GGGCGTCAAA AATAATTGT ATCGCACCTT AGAGGCTGTA

1521 CTCGGGTACG TGTTCCGGAC ATGGGCTCTT CTCCGGTAGC GGCAGGAGCTT CCACATCCGA GCCCTGGTCC CATCCGTCCA
GAGCCCATGC ACAAGGCCGTG TACCCGAGAA GAGGCCATCG CCGCTCGAA GGTGTAGGCT CGGGACCAGG GTAGGCAGGT

1601 GCGGCTCATG GTCGCTCGGC AGCTCCTTGC TCCTAACAGT GGAGGCCAGA CTTAGGCACA GCACAATGCC CACCACCACC
CGCCGAGTAC CAGCGAGCCG TCGAGGAACG AGGATTGTCA CCTCCGGTCT GAATCCGTGT CGTGTTACGG GTGGTGGTGG

1681 AGTGTGCCGC ACAAGGCCGT GCGGCTAGGG TATGTGTCTG AAAATGAGCT CGGAGATTGG GCTCGCACCT GGACGCAGAT
TCACACGGCG TGTTCCGGCA CCGCCATCCC ATACACAGAC TTTTACTCGA GCCTCTAACC CGAGCGTGA CCTGCGTCTA

1761 GGAAGACTTA AGGCAGCGGC AGAAGAAGAT GCAGGCAGCT GAGTTGTTGT ATTCTGATAA GAGTCAGAGG TAACTCCCGT
CCTTCTGAAT TCCGTCGCCG TCTTCTTCTA CGTCCGTCGA CTCACAAACA TAAGACTATT CTCAGTCTCC ATTGAGGGCA

1841 TCGGGTGCTG TTAACGGTGG AGGGCAGTGT AGTCTGAGCA GTACTCGTTG CTGCCGCGCG CGCCACCAGA CATAATAGCT
ACGCCACGAC AATTGCCACC TCCCGTCACA TCAGACTCGT CATGAGCAAC GACGGCGCGC GCGGTGGTCT GTATTATCGA

+2 M A A
EcoRI

1921 GACAGACTAA CAGACTGTTT CTTTCCATGG GTCTTTTCTG CAGTCACCGT CGTCGACCTA AGAATTCACC ATGGCTGCAT
CTGTCTGATT GTCTGACAAG GAAAGGTACC CAGAAAAGAC GTCAGTGGCA GCAGCTGGAT TCTTAAGTGG TACCGACGTA

+2 Y A A Q G Y K V L V L N P S V A A T L G F G A Y M S K
2001 ATGCAGCTCA GGGCTATAAG GTGCTAGTAC TCAACCCCTC TGTTGCTGCA ACACTGGGCT TTGGTGCTTA CATGTCCAAG
TACGTCGAGT CCCGATATTC CACGATCATG AGTTGGGGAG ACAACGACGT TGTGACCCGA AACCACGAAT GTACAGGTTT

+2 A H G I D P N I R T G V R T I T T G S P I T Y S T Y G
2081 GCTCATGGGA TCGATCCTAA CATCAGGACC GGGGTGAGAA CAATTACCAC TGGCAGCCCC ATCAGTACT CCACCTACGG
CGAGTACCCT AGCTAGGATT GTAGTCTG GCGGCTCTT GTTAATGGTG ACCGTCGGGG TAGTGCATGA GGTGGATGCC

+2 K F L A D G G C S G G A Y D I I I C D E C H S T D A
2161 CAAGTTCCTT GCCGACGGCG GGTGCTCGGG GGGCGCTTAT GACATAATAA TTTGTGACGA GTGCCACTCC ACGGATGCCA
GTTCAAGGAA CGGCTGCCGC CCACGAGCCC CCGCGAATA CTGTATTATT AAACACTGCT CACGGTGAGG TGCTACGGT

+2 T S I L G I G T V L D Q A E T A G A R L V V L A T A T
2241 CATCCATCTT GGGCATTGGC ACTGTCCTTG ACCAAGCAGA GACTGCGGGG GCGAGACTGG TTGTGCTCGC CACCGCCACC
GTAGGTAGAA CCCGTAACCG TGACAGGAAC TGTTCTGTCT CTGACGCCCC CGCTCTGACC AACACGAGCG GTGGCGGTGG

+2 P P G S V T V P H P N I E E V A L S T T G E I P F Y G
2321 CCTCCGGGCT CCGTCACTGT GCGCCATCCC AACATCGAGG AGGTTGCTCT GTCCACCACC GGAGAGATCC CTTTTACGG
GGAGGCCCGA GGCAGTGACA CCGGGTAGGG TTGTAGTCC TCCAACGAGA CAGGTGGTGG CCTCTCTAGG GAAAAATGCC

+2 K A I P L E V I K G G R H L I F C H S K K K C D E L
2401 CAAGGCTATC CCCCTCGAAG TAATCAAGGG GGGGAGACAT CTCATCTTCT GTCATTCAAA GAAGAAGTGC GACGAAGTGC
GTTCCGATAG GGGGAGCTTC ATTAGTTC CCGCTCTGTA GAGTAGAAGA CAGTAAGTTT CTGTTTACG CTGCTTGAGC

+2 A A K L V A L G I N A V A Y Y R G L D V S V I P T S G
2481 CCGCAAAGCT GGTGCGATTG GGCATCAATG CCGTGGCCTA CTACCGCGGT CTTGACGTGT CCGTCATCCC GACCAGCGGC
GGCGTTTCGA CCAGCGTAAC CCGTAGTTAC GGCACCGGAT GATGGCGCCA GAACTGCACA GGCAGTAGGG CTGGTGGCGG

+2 D V V V V A T D A L M T G Y T G D F D S V I D C N T C
2561 GATGTTGTCG TCGTGGAAC CGATGCCCTC ATGACCGGCT ATACCGGCGA CTTGACTCG GTGATAGACT GCAATACGTG
CTACAACAGC AGCACCCTTG GCTACGGGAG TACTGGCCGA TATGGCCGCT GAAGCTGAGC CACTATCTGA CGTTATGCAC

pCMV-NS35

FIGURE 3 - Page 3

+2 V T Q T V D F S L D P T F T I E T I T L P Q D A V S
 2641 . TGTCACCCAG ACAGTCGATT TCAGCCTTGA CCCTACCTTC ACCATTGAGA CAATCACGCT CCCCCAAGAT GCTGTCTCCC
 ACAGTGGGTC TGTCAGCTAA AGTCGGAAGT GGGATGGAAG TGGTAACTCT GTTAGTGCGA GGGGGTTCTA CGACAGAGGG

+2 R T Q R R G R T G R G K P G I Y R F V A P G E R P S G
 2721 GCACTCAACG TCGGGGAGG ACTGGCAGGG GGAAGCCAGG CATCTACAGA TTTGTGGCAC CGGGGGAGCG CCCCTCCGGC
 CGTGAGTTGC AGCCCCGTCC TGACCGTCCC CCTTCGGTCC GTAGATGTCT AAACACCGTG GCCCCCTCGC GGGGAGGCCG

+2 M F D S S V L C E C Y D A G C A W Y E L T P A E T T V
 2801 ATGTTCTGACT CGTCCGTCTT CTGTGAGTGC TATGACGCAG GCTGTGCTTG GTATGAGCTC ACGCCCGCCG AGACTACAGT
 TACAAGCTGA GCAGGCAGGA GAACTCACG ATACTGCGTC CGACACGAAC CATACTCGAG TGCGGGCGGC TCTGATGTCA

+2 R L R A Y M N T P G L P V C Q D H L E F W E G V F T
 2881 TAGGCTACGA GCGTACATGA ACACCCCGGG GCTTCCCGTG TGCCAGGACC ATCTTGAATT TTGGGAGGGC GTCTTTACAG
 ATCCGATGCT CGCATGTACT TGTGGGGCCC CGAAGGGCAC ACGGTCCTGG TAGAACTTAA AACCCCTCCG CAGAAATGTC
 StuI

+2 G L T H I D A H F L S Q T K Q S G E N L P Y L V A Y Q
 2961 GCCTCACTCA TATAGATGCC CACTTTCTAT CCCAGACAAA GCAGAGTGGG GAGAACCTTC CTTACCTGGT AGCGTACCAA
 CGGAGTGAGT ATATCTACGG GTGAAAGATA GGGTCTGTTT CGTCTCACCC CTCTTGAAG GAATGGACCA TCGCATGGTT

+2 A T V C A R A Q A P P P S W D Q M W K C L I R L K P T
 3041 GCCACCGTGT GCGCTAGGGC TCAAGCCCTT CCCCCATCGT GGGACCAGAT GTGGAAGTGT TTGATTGCGC TCAAGCCAC
 CGGTGGCACA CGCGATCCCG AGTTCGGGGA GGGGGTAGCA CCCTGGTCTA CACCTTCACA AACTAAGCGG AGTTCGGGTG

+2 L H G P T P L L Y R L G A V Q N E I T L T H P V T K
 3121 CCTCCATGGG CCAACACCCC TGCTATACAG ACTGGGCGCT GTTCAGAATG AAATCACCTT GACGCACCCA GTCACCAAAT
 GGAGTACCC GGTGTGGGG ACGATATGTC TGACCCGCGA CAAGTCTTAC TTTAGTGGGA CTGCGTGGGT CAGTGGTTTA

+2 Y I M T C M S A D L E V V T S T W V L V G G V L A A L
 3201 ACATCATGAC ATGCATGTCG GCCGACCTGG AGGTCGTCAC GAGCACCTGG GTGCTCGTTG GCGGCGTCTT GGCTGCTTTG
 TGTAGTACTG TACGTACAGC CGGCTGGACC TCCAGCAGTG CTCGTGGACC CACGAGCAAC CGCCGCGAGG CCGACGAAAC

+2 A A Y C L S T G C V V I V G R V V L S G K P A I I P D
 3281 GCCGCGTATT GCCTGTCAAC AGGCTGCGTG GTCATAGTGG GCAGGGTCTG CTTGTCCGGG AAGCCGGCAA TCATACCTGA
 CGGCGCATAA CGGACAGTTG TCCGACGCAC CAGTATCACC CGTCCCAGCA GAACAGGCCC TTCGGCCGTT AGTATGGACT

+2 R E V L Y R E F D E M E E C S Q H L P Y I E Q G M M
 3361 CAGGGAAGTC CTCTACCGAG AGTTCGATGA GATGGAAGAG TGCTCTCAGC ACTTACCGTA CATCGAGCAA GGGATGATGC
 GTCCTTTCAG GAGATGGCTC TCAAGCTACT CTACCTTCTC ACGAGAGTCG TGAATGGCAT GTAGCTCGTT CCTACTACG

+2 L A E Q F K Q K A L G L L Q T A S R Q A E V I A P A V
 3441 TCGCCGAGCA GTTCAAGCAG AAGGCCCTCG GCCTCCTGCA GACCGCGTCC CGTCAGGCAG AGGTTATCGC CCCTGCTGTC
 AGCGGCTCGT CAAGTTCGTC TTCCGGGAGC CGGAGGACGT CTGGCGCAGG GCAGTCCGTC TCCAATAGCG GGGACGACAG

+2 Q T N W Q K L E T F W A K H M W N F I S G I Q Y L A G
 3521 CAGACCAACT GGCAAAACT CGAGACCTTC TGGGCGAAGC ATATGTGGAA CTTATCATG GGGATACAAT ACTTGGCGGG
 GTCTGTTGA CCGTTTTTGA GCTCTGGAAG ACCCGCTTCG TATACACCTT GAAGTAGTCA CCCTATGTTA TGAACCGCCC

+2 L S T L P G N P A I A S L M A F T A A V T S P L T T
 3601 CTTGTCAACG CTGCCTGGTA ACCCGGCCAT TGCTTCATTG ATGGCTTTTA CAGCTGCTGT CACCAGCCCA CTAACCACTA
 GAACAGTTGC GACGGACCAT TGGGGCGGTA ACGAAGTAAC TACCGAAAAT GTCGACGACA GTGGTGGGT GATTGGTGT

+2 S Q T L L F N I L G G W V A A Q L A A P G A A T A F V
 3681 GCCAAACCTT CCTCTTCAAC ATATTGGGGG GGTGGGTGGC TGCCAGCTC GCCGCCCCCG GTGCCGCTAC TGCCTTTGTG
 CGGTTTGGGA GGAGAAGTTG TATAACCCCC CCACCCACCG ACGGGTCGAG CGGCGGGGGC CACGGCGATG ACGGAAACAC

FIGURE 3 - Page 4

+2 G A G L A G A A I G S V G L G K V L I D I L A G Y G A
 3761 G G C G T G G C T T A G C T G G C G C G C C A T C G G C A G T G T T G G A C T G G G A A G G T C C T C A T A G A C A T C C T T G C A G G G T A T G G C G A
 C C G C G A C C G A A T C G A C C G C G C G G T A G C C G T C A C A A C C T G A C C C T T C C A G G A T A T C T G T A G G A A C G T C C C A T A C C G C G

+2 G V A G A L V A F K I M S G E V P S T E D L V N L L
 3841 G G C G T G G C G G G A G C T C T T G T G C A T T C A A G A T C A T G A G C G G T G A G G T C C C T C C A C G G A G G A C C T G G T C A A T C T A C T G C
 C C C G C A C C G C C C T C G A G A A C A C C G T A A G T T C T A G T A C T C G C C A C T C C A G G G A G G T G C C T C T G G A C C A G T T A G A T G A C G

+2 P A I L S P G A L V V G V V C A A I L R R H V G P G E
 3921 C C G C C A T C C T C T C G C C C G G A G C C T C G T A G T C G G C G T G G T C T G T G C A G C A A T A C T G C G C C G G C A C G T T G G C C G G G C G A G
 G G C G T A G G A G A G C G G G C C T C G G G A G C A T C A G C C G C A C C A G A C A C G T C G T T A T G A C G C G G C C G T G C A A C C G G C C C G C T C

+2 G A V Q W M N R L I A F A S R G N H V S P T H Y V P E
 4001 G G G G C A G T G C A G T G G A T G A A C C G G C T G A T A G C T T C G C C T C C C G G G G G A A C C A T G T T T C C C C A C G C A C T A C G T G C C G G A
 C C C G T C A C G T C A C C T A C T T G G C C G A C T A T C G G A A G C G G A G G G C C C C T T G G T A C A A A G G G G T G C G T G A T G C A C G G C C T

+2 S D A A A R V T A I L S S L T V T Q L L R R L H Q W
 4081 G A G C G A T G C A G C T G C C C G C G T A C T G C C A T A C T C A G C A G C C T C A C T G T A A C C A G C T C C T G A G G C G A C T G C A C C A G T G G A
 C T C G C T A C G T C G A C G G G C G C A G T G A C G G T A T G A G T C G T C G G A G T G A C A T T G G T C G A G G A C T C C G C T G A C G T G G T C A C C T

+2 I S S E C T T P C S G S W L R D I W D W I C E V L S D
 4161 T A A G C T C G G A G T G T A C C A C T C C A T G C T C C G G T T C C T G G C T A A G G G A C A T C T G G G A C T G G A T A T G C G A G G T G T T G A G C G A C
 A T T C G A G C C T C A C A T G G T G A G G T A C G A G G C C A A G G A C C G A T T C C T G T A G A C C T G A C C T A T A C G C T C C A C A A C T C G C T G

+2 F K T W L K A K L M P Q L P G I P F V S C Q R G Y K G
 BamHI
 ~~~~~  
 4241 T T T A A G A C C T G G C T A A A A G C T A A G C T C A T G C C A C A G C T G C T G G G A T C C C T T T G T G T C C T G C C A G C G C G G G T A T A A G G G  
 A A A T T C T G G A C C G A T T T T C G A T T C G A G T A C G G T G T C G A C G G A C C C T A G G G G A A C A C A G G A C G G T C G C G C C A T A T T C C C

---

+2 V W R G D G I M H T R C H C G A E I T G H V K N G T  
 4321 G G T C T G G C G A G G G G A C G G C A T C A T G C A C A C T C G C T G C C A C T G T G G A G C T G A G A T C A C T G G A C A T G T C A A A A A C G G G G A C G A  
 C C A G A C C G C T C C C C T G C C G T A G T A C G T G T G A G C G A C G G T G A C A C C C T C G A C T C T A G T G A C C T G T A C A G T T T T G C C C T G C T

---

+2 M R I V G P R T C R N M W S G T F P I N A Y T T G P C  
 4401 T G A G G A T C G T C G G T C C T A G G A C C T G C A G G A A C A T G T G G A G T G G G A C C T T C C C A T T A A T G C C T A C A C C A C G G G C C C T G T  
 A C T C C T A G C A G C C A G G A T C C T G G A C G T C C T G T A C A C C T C A C C T G G A A G G G G T A A T T A C G G A T G T G G T G C C C G G G G A C A

---

+2 T P L P A P N Y T F A L W R V S A E E Y V E I R Q V G  
 4481 A C C C C C C T T C T G C G C C G A A C T A C A C G T T C G C G T A T G G A G G G T G T C T G C A G A G G A A T A C G T G G A G A T A A G G C A G G T G G G  
 T G G G G G A A G G A C G C G G C T T G A T G T G C A A G C G C G A T A C C T C C C A C A G A C G T C T C C T T A T G C A C C T C T A T T C C G T C C A C C C

---

+2 D F H Y V T G M T T D N L K C P C Q V P S P E F F T  
 4561 G G A C T T C C A C T A C T G A C G G G T A T G A C T A C T G A C A A T C T T A A T G C C C G T G C C A G G T C C C A T C G C C C G A A T T T T T C A C A G  
 C C T G A A G G T G A T G C A C T G C C A T A C T G A T G A C T G T T A G A A T T A C G G G C A C G G T C C A G G G T A G C G G G C T T A A A A A G T G T C

---

+2 E L D G V R L H R F A P P C K P L L R E E V S F R V G  
 4641 A A T T G G A C C G G G T G C G C C T A C A T A G G T T T G C G C C C C C T G C A A G C C C T T G C T G C G G G A G G A G G T A T C A T T C A G A G T A G G A  
 T T A A C C T G C C C A C G C G G A T G T A T C C A A A C G C G G G G G G A C G T T C G G G A A C G A C G C C C T C C T C C A T A G T A A G T C T A T C C T

---

+2 L H E Y P V G S Q L P C E P E P D V A V L T S M L T D  
 4721 C T C C A C G A A T A C C C G G T A G G G T C G C A A T T A C C T T G C G A G C C C G A A C C G T G C C G T G T T G A C G T C C A T G T G C T A C T G A  
 G A G G T G C T T A T G G G C C A T C C A G C G T T A A T G G A A C G C T C G G G C T T G G C C T G C A C C G G C A C A A C T G C A G G T A C G A G T G A C T

---

+2 P S H I T A E A A G R R L A R G S P P S V A S S S A  
 4801 T C C C T C C C A T A T A A C A G C A G A G G C G G C C G G G C G A A G G T T G C G A G G G G A T C A C C C C C C T C T G T G G C C A G C T C C T C G G C T A  
 A G G G A G G G T A T A T T G T C G T C T C C G C C G G C C G C T T C C A A C C G T C C C C T A G T G G G G G A G A C A C C G G T C G A G G A G C C G A T

## FIGURE 3 - Page 5

+2 S Q L S A P S L K A T C T A N H D S P D A E L I E A N  
 4881 GCCAGCTATC CGCTCCATCT CTAAGGCAA CTTGCACCGC TAACCATGAC TCCCCTGATG CTGAGCTCAT AGAGGCCAAC  
 CCGTTCGATAG GCGAGGTAGA GAGTTCGGTT GAACGTGGCG ATTGGTACTG AGGGGACTAC GACTCGAGTA TCTCCGGTTG

---

+2 L L W R Q E M G G N I T R V E S E N K V V I L D S F D  
 4961 CTCCTATGGA GGCAGGAGAT GGGCGGCAAC ATCACCAGGG TTGAGTCAGA AAACAAAGTG GTGATTCTGG ACTCCTTCGA  
 GAGGATACCT CCGTCCTCTA CCCGCCGTTG TAGTGGTCCC AACTCAGTCT TTTGTTTCAC CACTAAGACC TGAGGAAGCT

---

+2 P L V A E E D E R E I S V P A E I L R K S R R F A Q  
 5041 TCCGCTTGTG GCGGAGGAGG ACGAGCGGGA GATCTCCGTA CCCGCAGAAA TCCTGCGGAA GTCTCGGAGA TTCGCCCAGG  
 AGGCGAACAC CGCTCCTCC TGCTCGCCCT CTAGAGGCAT GGGCGTCTTT AGGACGCCTT CAGAGCCTCT AAGCGGGTCC

---

+2 A L P V W A R P D Y N P P L V E T W K K P D Y E P P V  
 5121 CCCTGCCCCG TTGGGCGCGG CCGGACTATA ACCCCCCGCT AGTGGAGACG TGGAAAAAGC CCGACTACGA ACCACCTGTG  
 GGGACGGGCA AACCCGCGCC GGCCTGATAT TGGGGGGCGA TCACCTCTGC ACCTTTTTTCG GGCTGATGCT TGGTGGACAC

---

+2 V H G C P L P P P K S P P V P P P R K K R T V V L T E  
 5201 GTCCATGGCT GCCCGCTTCC ACCTCCAAAG TCCCCTCCTG TGCCTCCGCC TCGGAAGAAG CGGACGGTGG TCCTCACTGA  
 CAGGTACCGA CGGGCGAAGG TGGAGGTTTC AGGGGAGGAC ACGGAGGCGG AGCCTTCTTC GCCTGCCACC AGGAGTGACT

---

+2 S T L S T A L A E L A T R S F G S S S T S G I T G D  
 5281 ATCAACCCTA TCTACTGCCT TGGCCGAGCT CGCCACCAGA AGCTTTGGCA GCTCCTCAAC TTCCGGCATT ACGGGCGACA  
 TAGTTGGGAT AGATGACGGA ACCGGCTCGA GCGGTGGTCT TCGAAACCGT CGAGGAGTTG AAGGCCGTAA TGCCCGCTGT

---

+2 N T T T S S E P A P S G C P P D S D A E S Y S S M P P  
 5361 ATACGACAAC ATCCTCTGAG CCCGCCCTT CTGGTGCCC CCCCAGTCC GACGCTGAGT CCTATTCTC CATGCCCCC  
 TATGCTGTTG TAGGAGACTC GGGCGGGGAA GACCGACGGG GGGGCTGAGG CTGCGACTCA GGATAAGGAG GTACGGGGG

---

+2 L E G E P G D P D L S D G S W S T V S S E A N A E D V  
 ~~~~~  
 BamHI

5441 CTGGAGGGGG AGCCTGGGGA TCCGATCTT AGCGACGGGT CATGGTCAAC GGTCAAGTAGT GAGGCCAACG CGGAGGATGT
 GACCTCCCCC TCGGACCCCT AGGCCTAGAA TCGTGCCCA GTACCAGTTG CCAGTCATCA CTCCGGTTGC GCCTCCTACA

+2 V C C S M S Y S W T G A L V T P C A A E E Q K L P I
 5521 CGTGTGCTGC TCAATGTCTT ACTCTGGAC AGGCGCACTC GTCACCCCGT GCGCCGCGGA AGAACAGAAA CTGCCCATCA
 GCACACGACG AGTTACAGAA TGAGAACCTG TCCGCGTGAG CAGTGGGGCA CGCGGCGCCT TCTTGTCTTT GACGGGTAGT

+2 N A L S N S L L R H H N L V Y S T T S R S A C Q R Q K
 5601 ATGCACTAAG CAACTCGTTG CTACGTCAAC ACAATTGGT GTATTCCACC ACCTCAGCA GTGCTTGCCA AAGGCAGAAG
 TACGTGATTC GTTGAGCAAC GATGCAGTGG TGTTAAACCA CATAAGGTGG TGGAGTGCCT CACGAACGGT TTCCGTCTTC

+2 K V T F D R L Q V L D S H Y Q D V L K E V K A A A S K
 5681 AAAGTCACAT TTGACAGACT GCAAGTTCTG GACAGCCATT ACCAGGACGT ACTCAAGGAG GTTAAAGCAG CGGCGTCAAA
 TTTCAAGTGA AACTGTCTGA CGTTCAAGAC CTGTGCGTAA TGGTCTGCA TGAGTTCCTC CAATTCGTC GCCGAGTTT

+2 V K A N L L S V E E A C S L T P P H S A K S K F G Y
 5761 AGTGAAGGCT AACTTGCTAT CCGTAGAGGA AGCTTGACGC CTGACGCCCC CACACTCAGC CAAATCCAAG TTTGGTTATG
 TCACTTCCGA TTGAACGATA GGCATCTCCT TCGAACGTCTG GACTGCGGGG GTGTGAGTCG GTTTAGGTTT AAACCAATAC

+2 G A K D V R C H A R K A V T H I N S V W K D L L E D N
 5841 GGGCAAAAGA CGTCCGTTGC CATGCCAGAA AGGCCGTAAC CCACATCAAC TCCGTGTGGA AAGACCTTCT GGAAGACAAT
 CCCGTTTTCT GCAGGCAACG GTACGCTCTT TCCGGCATTG GGTGTAGTTG AGGCACACCT TTCTGGAAGA CCTTCTGTGA

+2 V T P I D T T I M A K N E V F C V Q P E K G G R K P A
 5921 GTAACACCAA TAGACACTAC CATCATGGCT AAGAACGAGG TTTTCTGCGT TCAGCCTGAG AAGGGGGGTC GTAAGCCAGC
 CATTGTGGTT ATCTGTGATG GTAGTACCGA TTCTTGCTCC AAAAGACGCA AGTCGGACTC TTCCCCCAG CATTGCGTCC

FIGURE 3 - Page 6

+2 R L I V F P D L G V R V C E K M A L Y D V V T K L P
 6001 TCGTCTCATC GTGTTCCCG ATCTGGGCGT GCGCGTGTGC GAAAGATGG CTTGTACGA CGTGGTTACA AAGCTCCCT
 AGCAGAGTAG CACAAGGGG TAGACCCGCA CGCGCACACG CTTTCTACC GAAACATGCT GCACCAATGT TTCGAGGGGA

+2 L A V M G S S Y G F Q Y S P G Q R V E F L V Q A W K S
 EcoRI
 ~~~~~

6081 TGGCCGTGAT GGAAGCTCC TACGGATTCC AATACTCACC AGGACAGCGG GTTGAATTCC TCGTCAAGC GTGGAAGTCC  
 ACCGGCACTA CCCTTCGAGG ATGCCTAAGG TTATGAGTGG TCCTGTGCGC CAACTTAAGG AGCACGTTCC CACCTTCAGG

---

+2 K K T P M G F S Y D T R C F D S T V T E S D I R T E E  
 6161 AAGAAAACCC CAATGGGGT CTCGTATGAT ACCCGCTGCT TTGACTCCAC AGTCACTGAG AGCGACATCC GTACGGAGGA  
 TTCTTTTGGG GTTACCCCAA GAGCATACTA TGGGCGACGA AACTGAGGTG TCAGTGACTC TCGCTGTAGG CATGCCTCCT

---

+2 A I Y Q C C D L D P Q A R V A I K S L T E R L Y V G  
 6241 GGCAATCTAC CAATGTTGTG ACCTCGACCC CCAAGCCCGC GTGGCCATCA AGTCCCTCAC CGAGAGGCTT TATGTTGGGG  
 CCGTTAGATG GTTACAACAC TGGAGCTGGG GGTTCGGGCG CACCGGTAGT TCAGGGAGTG GCTCTCCGAA ATACAACCC

---

+2 G P L T N S R G E N C G Y R R C R A S G V L T T S C G  
 6321 GCCCTCTTAC CAATTCAAGG GGGGAGAAGT GCGGCTATCG CAGGTGCCGC GCGAGCGGCG TACTGACAACT TAGCTGTGGT  
 CGGGAGAATG GTTAAGTTCC CCCCTCTTGA CGCCGATAGC GTCCACGGCG CGCTCGCCGC ATGACTGTTG ATCGACACCA

---

+2 N T L T C Y I K A R A A C R A A G L Q D C T M L V C G  
 6401 AACACCCTCA CTTGCTACAT CAAGGCCCGG GCAGCCTGTC GAGCCGCGAG GCTCCAGGAC TGCACCATGC TCGTGTGTGG  
 TTGTGGGAGT GAACGATGTA GTTCCGGGCC CGTCGGACAG CTCGGCGTCC CGAGGTCTTG ACGTGGTACG AGCACACACC

---

+2 D D L V V I C E S A G V Q E D A A S L R A F T E A M  
 6481 CGACGACTTA GTCGTTATCT GTGAAAGCGC GGGGGTCCAG GAGGACGCGG CGAGCCTGAG AGCCTTCACG GAGGCTATGA  
 GCTGCTGAAT CAGCAATAGA CACTTTCGCG CCCCCAGGTC CTCCTGCGCC GCTCGGACTC TCGGAAGTGC CTCCGATACT

---

+2 T R Y S A P P G D P P Q P E Y D L E L I T S C S S N V  
 6561 CCAGTACTC CGCCCCCCT GGGGACCCCG CACAACCAGA ATACGACTTG GAGCTCATAA CATCATGCTC CTCCAACGTG  
 GTCCATGAG GCGGGGGGGA CCCCTGGGGG GTGTTGGTCT TATGCTGAAC CTCGAGTATT GTAGTACGAG GAGGTTGCAC

---

+2 S V A H D G A G K R V Y Y L T R D P T T P L A R A A W  
 6641 TCAGTCGCCC ACGACGGCGC TGGAAAGAGG GTCTACTACC TCACCCGTGA CCCTACAACC CCCCTCGCGA GAGGTGCGTG  
 AGTCAGCGGG TGCTGCCGCG ACCTTTCTCC CAGATGATGG AGTGGGCACT GGGATGTTGG GGGAGCGCT CTCGACGCAC

---

+2 E T A R H T P V N S W L G N I I M F A P T L W A R M  
 6721 GGAGACAGCA AGACACACTC CAGTCAATTC CTGGCTAGGC AACATAATCA TGTTTGCCCC CACACTGTGG GCGAGGATGA  
 CCTCTGTCGT TCTGTGTGAG GTCAGTTAAG GACCGATCCG TTGTATTAGT ACAAACGGGG GTGTGACACC CGTCTCTACT

---

+2 I L M T H F F S V L I A R D Q L E Q A L D C E I Y G A  
 6801 TACTGATGAC CCATTTCTTT AGCGTCCTTA TAGCCAGGGA CCAGCTTGAA CAGGCCCTCG ATTGCGAGAT CTACGGGGCC  
 ATGACTACTG GGTAAAGAAA TCGCAGGAAT ATCGGTCCCT GGTGCAACTT GTCCGGGAGC TAACGCTCTA GATGCCCCGG

---

+2 C Y S I E P L D L P P I I Q R L H G L S A F S L H S Y  
 6881 TGCTACTCCA TAGAACCCT GGATCTACCT CCAATCATTC AAAGACTCCA TGGCCTCAGC GCATTTTCAC TCCACAGTTA  
 ACGATGAGGT ATCTTGGTGA CCTAGATGGA GGTTAGTAAG TTTCTGAGGT ACCGGAGTCG CGTAAAAGTG AGGTGTCAAT

---

+2 S P G E I N R V A A C L R K L G V P P L R A W R H R  
 6961 CTCTCCAGGT GAAATCAATA GGTGGCCGCG ATGCCTCAGA AAATTGGGG TACCGCCCTT GCGAGCTTGG AGACACCGGG  
 GAGAGGTCCA CTTTAGTTAT CCCACCGGCG TACGGAGTCT TTTGAACCC ATGGCGGGAA CGCTCGAACC TCTGTGGCCC

---

+2 A R S V R A R L L A R G G R A A I C G K Y L F N W A V  
 7041 CCCGGAGCGT CCGCGCTAGG CTTCTGGCCA GAGGAGGCAG GGCTGCCATA TGTGGCAAGT ACCTCTTCAA CTGGGCAGTA  
 GGGCCTCGCA GCGCGATCC GAAGACCGGT CTCCTCCGTC CCGACGGTAT ACACCGTTCA TGGAGAAGTT GACCCGTCAT

0022445-43300



## FIGURE 3 - Page 7

+2 R T K L K L T P I A A A G Q L D L S G W F T A G Y S G  
 7121 AGAACAAAGC TCAAACTCAC TCCAATAGCG GCCGCTGGCC AGCTGGACTT GTCCGGCTGG TTCACGGCTG GCTACAGCGG  
 TCTTGTTCG AGTTTGAGTG AGGTTATCGC CGGCGACCGG TCGACCTGAA CAGGCCGACC AAGTGCCGAC CGATGTCGCG

---

+2 G D I Y H S V S H A R P R W I W F C L L L L A A G V  
 7201 GGGAGACATT TATCACAGCG TGTCTCATGC CCGGCCCGCG TGGATCTGGT TTTGCCTACT CCTGCTTGCT GCAGGGGTAG  
 CCCTCTGTAA ATAGTGTGCG ACAGAGTACG GGCCGGGGCG ACCTAGACCA AAACGGATGA GGACGAACGA CGTCCCCATC

---

+2 G I Y L L P N R  
 7281 GCATCTACCT CCTCCCCAAC CGATGAAGGT TGGGGTAAAC ACTCCGGCCT AAAAAAAAAA AAAAATCTAG AAAGGCGCGC  
 CGTAGATGGA GGAGGGGTG GCTACTTCCA ACCCCATTG TGAGGCCGGA TTTTTTTTTT TTTTATAGT TTTCCGCGCG

---

BamHI MluI  
 7361 CAAGATATCA AGGATCCACT ACGCGTTAGA GCTCGCTGAT CAGCCTCGAC TGTGCCTTCT AGTTGCCAGC CATCTGTTGT  
 GTTCTATAGT TCCTAGGTGA TCGCGAATCT CGAGCGACTA GTCGGAGCTG ACACGGAAGA TCAACGGTCG GTAGACAACA

---

7441 TTGCCCTCC CCGTGCCTT CTTGACCCT GGAAGGTGCC ACTCCCACTG TCCTTTCCTA ATAAAATGAG GAAATTGCAT  
 AACGGGGAGG GGGCACGGAA GGAAGTGGGA CCTTCCACGG TGAGGGTGAC AGGAAAGGAT TATTTTACTC CTTTAACTGA

---

7521 CGCATTGTCT GAGTAGGTGT CATTCTATTC TGGGGGGTGG GGTGGGGCAG GACAGCAAGG GGGAGGATTG GGAAGACAAT  
 GCGTAACAGA CTATCCACA GTAAGATAAG ACCCCCCACC CCACCCCGTC CTGTGCTTCC CCCTCCTAAC CTTTCTGTTA

---

7601 AGCAGGCATG CTGGGGAGCT CTTCCGCTTC CTCGCTCACT GACTCGCTGC GCTCGGTCGT TCGGCTGCGG CGAGCGGTAT  
 TCGTCCGTAC GACCCCTCGA GAAGGCGAAG GAGCGAGTGA CTGAGCGACG CGAGCCAGCA AGCCGACGCC GCTCGCCATA

---

7681 CAGCTCACTC AAAGGCGGTA ATACGGTTAT CCACAGAATC AGGGGATAAC GCAGGAAAGA ACATGTGAGC AAAAGGCCAG  
 GTCGAGTGAG TTTCCGCCAT TATGCCAATA GGTGTCTTAG TCCCCTATTG CGTCTTTTCT TGTACACTCG TTTTCCGGTC

---

7761 CAAAAGGCCA GGAACCGTAA AAAGGCCGCG TTGCTGGCGT TTTTCCATAG GCTCCGCCCC CCTGACGAGC ATCACAAAAA  
 GTTTTCCGGT CTTTGGCATT TTTCCGGCGC AACGACCGCA AAAAGGTATC CGAGGCGGGG GGACTGCTCG TAGTGTTTTT

---

7841 TCGACGCTCA AGTCAGAGGT GGCGAAACCC GACAGGACTA TAAAGATAAC AGGCGTTTCC CCCTGGAAGC TCCCTCGTGC  
 AGCTGCGAGT TCAGTCTCCA CCGCTTTGGG CTGTCTGATG ATTTCTATGG TCCGCAAAGG GGGACCTTCG AGGGAGCAGC

---

7921 GCTCTCCTGT TCCGACCCTG CCGCTTACCG GATACCTGTC CGCCTTTCTC CCTTCGGGAA GCGTGGCGCT TTCTCAATGC  
 CGAGAGGACA AGGCTGGGAC GGCGAATGGC CTATGGACAG GCGGAAAGAG GGAAGCCCTT CGCACC CGA AAGAGTTACG

---

8001 TCACGCTGTA GGTATCTCAG TTCGGTGTAG GTCGTTGCT CCAAGCTGGG CTGTGTGCAC GAACCCCCCG TTCAGCCCGA  
 AGTGCGACAT CCATAGAGTC AAGCCACATC CAGCAAGCGA GGTTCGACCC GACACACGTG CTTGGGGGGC AAGTCGGGCT

---

8081 CCGCTGCGCC TTATCCGGTA ACTATCGTCT TGAGTCCAAC CCGGTAAGAC ACGACTTATC GCCACTGGCA GCAGCCACTG  
 GGCGACGCGG AATAGGCCAT TGATAGCAGA ACTCAGGTG GGCCATTCTG TGCTGAATAG CGGTGACCGT CGTCGGTGAC

---

8161 GTAACAGGAT TAGCAGAGCG AGGTATGTAG GCGGTGCTAC AGAGTTCTTG AAGTGGTGGC CTAACACGG CTACACTAGA  
 CATTGTCTTA ATCGTCTCGC TCCATACATC CGCCACGATG TCTCAAGAAC TTCACCACCG GATTGATGCC GATGTGATCT

---

8241 AGGACAGTAT TTGGTATCTG CGCTCTGCTG AAGCCAGTTA CCTTCGGAAA AAGAGTTGGT AGCTCTTGAT CCGGCAACA  
 TCCTGTCATA AACCATAGAC GCGAGACGAC TTCGGTCAAT GGAAGCCTTT TTCTCAACCA TCGAGAACCTA GGCCGTTTGT

---

8321 AACCACCGCT GGTAGCGGTG GTTTTTTTGT TTGCAAGCAG CAGATTACGC GCAGAAAAAA AGGATCTCAA GAAGATCCTT  
 TTGGTGGCGA CCATCGCCAC CAAAAAACA AACGTTGCTG GTCTAATGCG CGTCTTTTTT TCCTAGAGTT CTTCTAGGAA

---

8401 TGATCTTTTC TACGGGGTCT GACGCTCAGT GGAACGAAAA CTCACGTTAA GGGATTTTGG TCATGAGATT ATCAAAAAGG  
 ACTAGAAAAG ATGCCCCAGA CTGCGAGTCA CCTTGCTTTT GAGTGCAATT CCCTAAAACC AGTACTCTAA TAGTTTTTCC

00327 6472460

## FIGURE 3 - Page 8

8481 ATCTTCACCT AGATCCTTTT AAATTA AAAA TGAAGTTT AATCAATCTA AAGTATATAT GAGTAAACTT GGTCTGACAG  
TAGAAGTGGA TCTAGGAAAA TTTAATTTTT ACTTCAAAAT TTAGTTAGAT TTCATATATA CTCATTTGAA CCAGACTGTC

---

8561 TTACCAATGC TTAATCAGTG AGGCACCTAT CTCAGCGATC TGTCTATTTT GTTCATCCAT AGTTGCCTGA CTCCCCGTGG  
AATGGTTACG AATTAGTCAC TCCGTGGATA GAGTCGCTAG ACAGATAAAG CAAGTAGGTA TCAACGGACT GAGGGGCGAG

---

8641 TGTAGATAAC TACGATACGG GAGGGCTTAC CATCTGGCCC CAGTGCTGCA ATGATACCGC GAGACCCACG CTCACCGGCT  
ACATCTATTG ATGCTATGCC CTCCCGAATG GTAGACGGGG GTCACGACGT TACTATGGCG CTCTGGGTGC GAGTGGCCGA

---

8721 CCAGATTTAT CAGCAATAAA CCAGCCAGCC GGAAGGGCCG AGCGCAGAAG TGGTCCTGCA ACTTTATCCG CCTCCATCCA  
GGTCTAAATA GTCGTTATTT GGTGGGTGCG CCTTCCCGGC TCGCGTCTTC ACCAGGACGT TGAAATAGGC GGAGGTAGGT

---

8801 GTCTATTAAT TGTGCGGGG AAGCTAGAGT AAGTAGTTCG CCAGTTAATA GTTTGCGCAA CGTTGTTGCC ATTGCTACAG  
CAGATAATTA ACAACGGCCC TTCGATCTCA TTCATCAAGC GGTCAATTAT CAAACGCGTT GCAACAACGG TAACGATGTC

---

8881 GCATCGTGGT GTCACGCTCG TCGTTTGGTA TGGCTTCATT CAGCTCCGGT TCCCAACGAT CAAGGCGAGT TACATGATCC  
CGTAGCACCA CAGTGCGAGC AGCAAACCAT ACCGAAGTAA GTCGAGGCCA AGGGTTGCTA GTTCCGCTCA ATGTACTAGG

---

8961 CCCATGTTGT GCAAAAAAGC GGTTAGCTCC TTCGGTCTCT CGATCGTTGT CAGAAGTAAG TTGGCCGCGAG TGTATCACT  
GGGTACAACA CGTTTTTCG CCAATCGAGG AAGCCAGGAG GCTAGCAACA GTCTTCATTC AACC GGCGTC ACAATAGTGA

---

9041 CATGGTTATG GCAGCACTGC ATAATTCTCT TACTGTCTAT CCATCCGTAA GATGCTTTTC TGTGACTGGT GAGTACTCAA  
GTACCAATAC CGTCGTGACG TATTAAGAGA ATGACAGTAC GGTAGGCATT CTACGAAAAG ACACTGACCA CTCATGAGTT

---

9121 CCAAGTCATT CTGAGAATAG TGTATGCGGC GACCGAGTTG CTCTTGCCCG GCGTCAATAC GGGATAATAC CGCGCCACAT  
GGTTCACTAA GACTCTTATC ACATACGCCG CTGGCTCAAC GAGAACGGGG CGCAGTTATG CCCTATTATG GCGCGGTGTA

---

9201 AGCAGAACTT TAAAGTGCT CATCATTGGA AAACGTTCTT CGGGGCGAAA ACTCTCAAGG ATCTTACCGC TGTGAGATC  
TCGTCTTGAA ATTTTCACGA GTAGTAACCT TTTGCAAGAA GCCCGGCTTT TGAGAGTTCC TAGAATGGCG ACAACTCTAG

---

9281 CAGTTCGATG TAACCCACTC GTGCACCCAA CTGATCTTCA GCATCTTTTA CTTTCACCAG CGTTTCTGGG TGAGCAAAAA  
GTCAAGCTAC ATTGGGTGAG CACGTGGGTT GACTAGAAGT CGTAGAAAAT GAAAGTGGTC GCAAAGACCC ACTCGTTTTT

---

9361 CAGGAAGGCA AAATGCCGCA AAAAAGGGAA TAAGGGCGAC ACGGAAATGT TGAATACTCA TACTCTTCCT TTTTCAATAT  
GTCCTTCGCT TTTACGGCGT TTTTCCCTT ATTCCCGCTG TGCCTTTACA ACTTATGAGT ATGAGAAGGA AAAAGTTATA

---

9441 TATTGAAGCA TTTATCAGGG TTATTGTCTC ATGAGCGGAT ACATATTTGA ATGTATTTAG AAAAATAAAC AAATAGGGGT  
ATAACTTCGT AAATAGTCCC AATAACAGAG TACTCGCCTA TGTATAAACT TACATAAATC TTTTATTG TTTATCCCCA

---

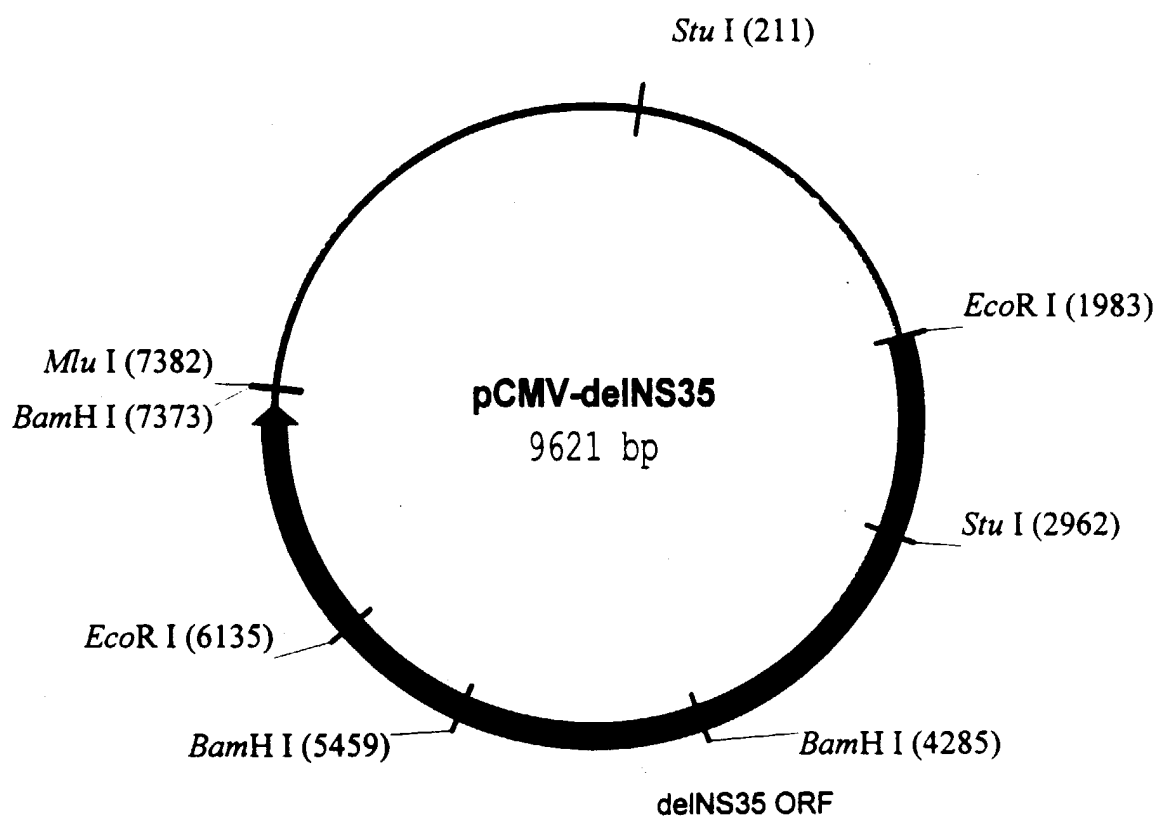
9521 TCCGCGCACA TTTCCCCGAA AAGTGCCACC TGACGTCTAA GAAACCATTA TTATCATGAC ATTAACCTAT AAAAATAGGC  
AGGCGCGTGT AAAGGGGCTT TTCACGGTGG ACTGCAGATT CTTTGTAAT AATAGTACTG TAATTGGATA TTTTATCCG

---

9601 GTATCACGAG GCCCTTTCGT C  
CATAGTGCTC CGGGAAAGCA G

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

**FIGURE 4**



|      |            |            |             |            |            |            |             |             |
|------|------------|------------|-------------|------------|------------|------------|-------------|-------------|
| 1    | TCGCGCGTTT | CGGTGATGAC | GGTGAAGAAC  | TCTGACACAT | GCAGCTCCCG | GAGACGGTCA | CAGCTTGTCT  | GTAAAGCGGAT |
|      | AGCGCGCAAA | GCCACTACTG | CCACTTTTGG  | AGACTGTGTA | CGTCGAGGGC | CTCTGCCAGT | GTCGAACAGA  | CATTGCGCTA  |
| 81   | GCCGGGAGCA | GACAAGCCCG | TCAGGGCGCG  | TCAGCGGGTG | TTGGCGGGTG | TCGGGGCTGG | CTTAACATATG | CGGCATCAGA  |
|      | CGGCCCTCGT | CTGTTTCGGG | AGTCCCGCGC  | AGTCGCCAC  | AACCGCCAC  | AGCCCCGACC | GAATTGATAC  | GCCGTAGTCT  |
|      |            |            |             |            | StuI       |            |             |             |
| 161  | GCAGATTGTA | CTGAGAGTGC | ACCATATGAA  | GCTTTTGTGA | AAAGCCTAGG | CCTCCAAAAA | AGCCTCCTCA  | CTACTTCTGG  |
|      | CGTCTAACAT | GACTCTCAGG | TGGTATACTT  | CGAAAAACGT | TTTCGGATCC | GGAGGTTTTT | TCGGAGGAGT  | GATGAAGACC  |
| 241  | AATAGCTCAG | AGGCCGAGGC | GGCCTCGGCC  | TCTGCATAAA | TAAAAAAAT  | TAGTCAGCCA | TGGGGCGGAG  | AATGGGCGGA  |
|      | TTATCGAGTC | TCCGGCTCCG | CCGGAGCCGG  | AGACGTATTT | ATTTTTTTTA | ATCAGTCGGT | ACCCCGCCTC  | TTACCCGCTC  |
| 321  | ACTGGGCGGG | GAGGGAATTA | TTGCTATTG   | GCCATTGCAT | ACGTTGTATC | TATATCATAA | TATGTACATT  | TATATTGGCT  |
|      | TGACCCGCCC | CTCCCTTAAT | AACCGATAAC  | CGGTAACGTA | TGCAACATAG | ATATAGTATT | ATACATGTAA  | ATATAACCGA  |
| 401  | CATGTCCAAT | ATGACCGCCA | TGTTGACATT  | GATTATTGAC | TAGTTATTAA | TAGTAATCAA | TTACGGGGTC  | ATTAGTTCAT  |
|      | GTACAGGTTA | TACTGGCGGT | ACAACGTAA   | CTAATAACTG | ATCAATAATT | ATCATTAGTT | AATGCCCCAG  | TAATCAAGTA  |
| 481  | AGCCCATATA | TGGAGTCCG  | CGTTACATAA  | CTTACGGTAA | ATGGCCCGCC | TGGCTGACCG | CCCAACGACC  | CCCGCCCAT   |
|      | TCGGGTATAT | ACCTCAAGGC | GCAATGTATT  | GAATGCCATT | TACCGGGCGG | ACCGACTGGC | GGGTTGCTGG  | GGCGGGGTAA  |
| 561  | GACGTCAATA | ATGACGTATG | TTCCCATAGT  | AACGCCAATA | GGGACTTTCC | ATTGACGTCA | ATGGGTGGAG  | TATTTACGGT  |
|      | CTGCAGTTAT | TACTGCATAC | AAGGGTATCA  | TTGCGGTTAT | CCCTGAAAGG | TAAGTGCAGT | TACCCACCTC  | ATAAATGCCA  |
| 641  | AAACTGCCCA | CTTGGCAGTA | CATCAAGTGT  | ATCATATGCC | AAGTCCGCCC | CCTATTGACG | TCAATGACGG  | TAAATGGCCC  |
|      | TTTGACGGGT | GAACCGTCAT | GTAGTTTACA  | TAGTATACGG | TTCAGGCGGG | GGATAACTGC | AGTTACTGCC  | ATTTACGGGG  |
| 721  | GCCTGGCATT | ATGCCCAGTA | CATGACCTTA  | CGGGACTTTC | CTACTTGGCA | GTACATCTAC | GTATTAGTCA  | TCGCTATTAC  |
|      | CGGACCGTAA | TACGGGTCAT | GTACTGGAAT  | GCCCTGAAAG | GATGAACCGT | CATGTAGATG | CATAATCAGT  | AGCGATAATG  |
| 801  | CATGGTGATG | CGGTTTTGGC | AGTACACCAA  | TGGGCGTGGA | TAGCGGTTTG | ACTCACGGGG | ATTTCCAAGT  | CTCCACCCCA  |
|      | GTACCACTAC | GCCAAAACCG | TCATGTGGTT  | ACCCGCACCT | ATCGCCAAAC | TGAGTGCCCC | TAAAGGTTCA  | GAGGTGGGGT  |
| 881  | TTGACGTCAA | TGGGAGTTTG | TTTTGGCACC  | AAAATCAACG | GGACTTTCCA | AAATGTCGTA | ATAACCCCGC  | CCCGTTGACG  |
|      | AACTGCAGTT | ACCTCAAAC  | AAAACCGTGG  | TTTTAGTTGC | CCTGAAAGGT | TTTACAGCAT | TATTGGGGCG  | GGGCAACTGC  |
| 961  | CAAATGGGCG | GTAGGCGTGT | ACGGTGGGAG  | GTCTATATAA | GCAGAGCTCG | TTTAGTGAAC | CGTCAGATCG  | CCTGGAGACG  |
|      | GTTTACCCGC | CATCCGCACA | TGCCACCCTC  | CAGATATATT | CGTCTCGAGC | AAATCACTTG | GCAGTCTAGC  | GGACCTCTGC  |
| 1041 | CCATCCACGC | TGTTTTGACC | TCCATAGAAG  | ACACCGGGAC | CGATCCAGCC | TCCGCGGGCG | GGAACGGTGC  | ATTGGAACGC  |
|      | GGTAGGTGCG | ACAAAACCTG | AGGTATCTTC  | TGTGGCCCTG | GCTAGGTGCG | AGGCGCCGGC | CCTTGCCACG  | TAACCTTGCG  |
| 1121 | GGATTCCCCG | TGCCAAGAGT | GACGTAAGTA  | CCGCCTATAG | ACTCTATAGG | CACACCCCTT | TGGCTCTTAT  | GCATGCTATA  |
|      | CCTAAGGGGC | ACGGTTCTCA | CTGCATTTCAT | GGCGGATATC | TGAGATATCC | GTGTGGGGAA | ACCGAGAATA  | CGTACGATAT  |
| 1201 | CTGTTTTTGG | CTTGGGGCCT | ATACACCCCC  | GCTCCTTATG | CTATAGGTGA | TGGTATAGCT | TAGCCTATAG  | GTGTGGGGTA  |
|      | GACAAAAACC | GAACCCCGGA | TATGTGGGGG  | CGAGGAATAC | GATATCCACT | ACCATATCGA | ATCGGATATC  | CACACCCAAT  |
| 1281 | TTGACCATTA | TTGACCACTC | CCCTATTGGT  | GACGATACTT | TCCATTACTA | ATCCATAACA | TGGCTCTTTG  | CCACAACAT   |
|      | AACTGGTAAT | AACTGGTGAG | GGGATAACCA  | CTGCTATGAA | AGGTAATGAT | TAGGTATTGT | ACCGAGAAAC  | GGTGTGATA   |
| 1361 | CTCTATTGGC | TATATGCCAA | TACTCTGTCC  | TTCAGAGACT | GACACGGACT | CTGTATTTTT | ACAGGATGGG  | GTCCATTTAT  |
|      | GAGATAACCG | ATATACGGTT | ATGAGACAGG  | AAGTCTCTGA | CTGTGCCTGA | GACATAAAAA | TGTCCTACCC  | CAGGTAAATA  |

## FIGURE 5 - Page 2

1441 TATTTACAAA TTCACATATA CAACAACGCC GTCCCCCGTG CCCGCAGTTT TTATTAAACA TAGCGTGGA TCTCCGACAT  
ATAAATGTTT AAGTGTATAT GTTGTTCGGG CAGGGGGCAC GGGCGTCAAA AATAATTTGT ATCGCACCT AGAGGCTGTA

---

1521 CTCGGGTACG TGTTCGGAC ATGGGCTCTT CTCCGGTAGC GCGGAGCTT CCACATCCGA GCCCTGGTCC CATCCGTTCA  
GAGCCCATGC ACAAGGCCTG TACCCGAGAA GAGGCCATCG CCGCCTCGAA GGTGTAGGCT CGGGACCAGG GTAGGCAGGT

---

1601 GCGGCTCATG GTCGCTCGGC AGCTCCTTGC TCCTAACAGT GGAGGCCAGA CTTAGGCACA GCACAATGCC CACCACCACC  
CGCCGAGTAC CAGCGAGCCG TCGAGGAACG AGGATTGTCA CCTCCGGTCT GAATCCGTGT CGTGTACGG GTGGTGGTGG

---

1681 AGTGTGCCGC ACAAGGCCGT GCGGCTAGGG TATGTGTCTG AAAATGAGCT CGGAGATTGG GCTCGCACCT GGACGCAGAT  
TCACACGGCG TGTTCGGCA CCGCCATCCC ATACACAGAC TTTTACTCGA GCCTCTAACC CGAGCGTGA CCTGCGTCTA

---

1761 GGAAGACTTA AGGCAGCGGC AGAAGAAGAT GCAGGCAGCT GAGTTGTTGT ATTCTGATAA GAGTCAGAGG TAACTCCCGT  
CCTTCTGAAT TCCGTCGCCG TCTTCTTCTA CGTCCGTCGA CTCAACAACA TAAGACTATT CTCAGTCTCC ATTGAGGGCA

---

1841 TGCGGTGCTG TTAACGGTGG AGGGCAGTGT AGTCTGAGCA GTACTCGTTG CTGCCGCGCG CGCCACCAGA CATAATAGCT  
ACGCCACGAC AATTGCCACC TCCCGTCACA TCAGACTCGT CATGAGCAAC GACGGCGCGC GCGGTGGTCT GTATTATCGA

---

+2 M A A  
EcoRI  
~~~~~

1921 GACAGACTAA CAGACTGTTT CTTTCCATGG GTCTTTTCTG CAGTCACCGT CGTCGACCTA AGAATTCACC ATGGCTGCAT
CTGTCTGATT GTCTGACAAG GAAAGGTACC CAGAAAAGAC GTCAGTGCCA GCAGCTGGAT TCTTAAGTGG TACCGACGTA

+2 Y A A Q G Y K V L V L N P S V A A T L G F G A Y M S K

2001 ATGCAGCTCA GGGCTATAAG GTGCTAGTAC TCAACCCCTC TGTTGCTGCA AACTGGGCT TTGGTGCTTA CATGTCCAAG
TACGTCGAGT CCCGATATTC CACGATCATG AGTTGGGGAG ACAACGACGT TGTGACCCGA AACCACGAAT GTACAGGTTT

+2 A H G I D P N I R T G V R T I T T G S P I T Y S T Y G

2081 GCTCATGGGA TCGATCCTAA CATCAGGACC GGGGTGAGAA CAATTACCAC TGGCAGCCCC ATCAGCTACT CCACCTACGG
CGAGTACCCT AGCTAGGATT GTAGTCTTGG CCCCCTCTT GTTAATGGTG ACCGTCGGGG TAGTGCATGA GGTGGATGCC

+2 K F L A D G G C S G G A Y D I I I C D E C H S T D A

2161 CAAGTTCCTT GCCGACGGCG GGTGCTCGGG GGGCGCTTAT GACATAATAA TTTGTGACGA GTGCCACTCC ACGGATGCCA
GTTCAAGGAA CGGCTGCCGC CCACGAGCCC CCCGCGAATA CTGTATTATT AAACACTGCT CACGGTGAGG TGCTACGGT

+2 T S I L G I G T V L D Q A E T A G A R L V V L A T A T

2241 CATCATCTT GGGCATTGGC ACTGTCCTTG ACCAAGCAGA GACTGCGGGG GCGAGACTGG TTGTGCTCGC CACCGCCACC
GTAGGTAGAA CCCGTAACCG TGACAGGAAC TGGTTCGTCT CTGACGCCCC CGCTCTGACC AACACGAGCG GTGGCGGTGG

+2 P P G S V T V P H P N I E E V A L S T T G E I P F Y G

2321 CCTCCGGGCT CCGTCACTGT GCGCCATCCC AACATCGAGG AGGTTGCTCT GTCCACCACC GGAGAGATCC CTTTTACGG
GGAGGCCCCG GGCAGTGACA CGGGGTAGGG TTGTAGTCTC TCCAACGAGA CAGGTGGTGG CCTCTCTAGG GAAAAATGCC

+2 K A I P L E V I K G G R H L I F C H S K K K C D E L

2401 CAAGGCTATC CCCCTCGAAG TAATCAAGGG GGGGAGACAT CTCATCTTCT GTCATTCAAA GAAGAAGTGC GACGAAGTGC
GTTCCGATAG GGGGAGCTTC ATTAGTTCCC CCCCTCTGTA GAGTAGAAGA CAGTAAGTTT CTTCTTCACG CTGCTTGAGC

+2 A A K L V A L G I N A V A Y Y R G L D V S V I P T S G

2481 CCGCAAAGCT GGTGCGATTG GGCATCAATG CCGTGGCCTA CTACCGCGGT CTTGACGTGT CCGTCATCCC GACCAGCGGC
GGCGTTTCGA CCAGCGTAAC CCGTAGTTAC GGCACCGGAT GATGGCGCCA GAACTGCACA GGCAGTAGGG CTGGTCCGCC

+2 D V V V V A T D A L M T G Y T G D F D S V I D C N T C

2561 GATGTTGTCG TCGTGGAAC CGATGCCCTC ATGACCGGCT ATACCGGCGA CTTGACTCG GTGATAGACT GCAATACGTG
CTACAACAGC AGCACCCTTG GCTACGGGAG TACTGGCCGA TATGGCCGCT GAAGCTGAGC CACTATCTGA CGTTATGCAC

0037624360

FIGURE 5 - Page 3

+2 V T Q T V D F S L D P T F T I E T I T L P Q D A V S
 2641 TGTCACCCAG ACAGTCGATT TCAGCCTTGA CCCTACCTTC ACCATTGAGA CAATCACGCT CCCCCAAGAT GCTGTCTCCC
 ACAGTGGGTC TGTCAGCTAA AGTCGGAAC GGGATGGAAG TGGTAACTCT GTTAGTGCGA GGGGGTTCTA CGACAGAGGG

+2 R T Q R R G R T G R G K P G I Y R F V A P G E R P S G
 2721 GCACTCAACG TCGGGGCAGG ACTGGCAGGG GGAAGCCAGG CATCTACAGA TTTGTGGCAC CGGGGGAGCG CCCCTCCGCG
 CGTGAGTTGC AGCCCCGTCC TGACCGTCCC CCTTCGGTCC GTAGATGTCT AAACACCGTG GCCCCCTCGC GGGGAGGCGG

+2 M F D S S V L C E C Y D A G C A W Y E L T P A E T T V
 2801 ATGTTCTGACT CGTCCGTCCT CTGTGAGTGC TATGACGCAG GCTGTGCTTG GTATGAGCTC ACGCCCGCCG AGACTACAGT
 TACAAGCTGA GCAGGCAGGA GACACTCAGC ATACTGCGTC CGACACGAAC CATACTCGAG TCGGGGCGGC TCTGATGTCA

+2 R L R A Y M N T P G L P V C Q D H L E F W E G V F T
 2881 TAGGCTACGA GCGTACATGA ACACCCCGGG GCTTCCCGTG TGCCAGGACC ATCTTGAATT TTGGGAGGGC GTCTTTACAG
 ATCCGATGCT CGCATGTACT TGTGGGGCCC CGAAGGGCAC ACGGTCTCTG TAGAACTTAA AACCTCCCCG CAGAAATGTC

+2 G L T H I D A H F L S Q T K Q S G E N L P Y L V A Y Q
 2961 GCCTCACTCA TATAGATGCC CACTTTCTAT CCCAGACAAA GCAGAGTGGG GAGAACCTTC CTTACCTGGT AGCGTACCAA
 CGGAGTGAGT ATATCTACGG GTGAAAGATA GGGTCTGTTT CGTCTCACCC CTCTTGAAG GAATGGACCA TCGCATGTTT

+2 A T V C A R A Q A P P P S W D Q M W K C L I R L K P T
 3041 GCCACCGTGT GCGCTAGGGC TCAAGCCCCT CCCCCATCGT GGGACCAGAT GTGGAAGTGT TTGATTGCGC TCAAGCCCAC
 CGGTGGCACA CGCGATCCCG AGTTCGGGGA GGGGGTAGCA CCCTGGTCTA CACCTTCACA AACTAAGCGG AGTTCGGGTG

+2 L H G P T P L L Y R L G A V Q N E I T L T H P V T K
 3121 CCTCCATGGG CCAACACCCC TGCTATACAG ACTGGGCGCT GTTCAGAATG AAATCACCTC GACGCACCCA GTCACCAAA
 GGAGGTACCC GGTGTGGGG ACGATATGTC TGACCCGCGA CAAGTCTTAC TTTAGTGGGA CTGCGTGGGT CAGTGGTTTA

+2 Y I M T C M S A D L E V V T S T W V L V G G V L A A L
 3201 ACATCATGAC ATGCATGTCG GCCGACCTGG AGGTCGTCAC GAGCACCTGG GTGCTCGTTG GCGGCGTCTT GGCTGCTTTG
 TGTAAGTACTG TACGTACAGC CGGCTGGACC TCCAGCAGTG CTCGTGGACC CACGAGCAAC CGCCGACGGA CCGACGAAAC

+2 A A Y C L S T G C V V I V G R V V L S G K P A I I P D
 3281 GCCGCGTATT GCCTGTCAAC AGGCTGCGTG GTCATAGTGG GCAGGTCGT CTTGTCCGGG AAGCCGGCAA TCATACCTGA
 CGGCGCATAA CGGACAGTTG TCCGACGCAC CAGTATCACC CGTCCCAGCA GAACAGGCCC TTCGGCCGTT AGTATGGACT

+2 R E V L Y R E F D E M E E C S Q H L P Y I E Q G M M
 3361 CAGGGAAGTC CTCTACCGAG AGTTCGATGA GATGGAAGAG TGCTCTCAGC ACTTACCGTA CATCGAGCAA GGGATGATGC
 GTCCCTTCAG GAGATGGCTC TCAAGCTACT CTACCTTCTC ACGAGAGTCG TGAATGGCAT GTAGCTCGTT CCCTACTACG

+2 L A E Q F K Q K A L G L L Q T A S R Q A E V I A P A V
 3441 TCGCCGAGCA GTTCAAGCAG AAGGCCCTCG GCCTCCTGCA GACCGCGTCC CGTCAGGCAG AGGTTATCGC CCCTGCTGTC
 AGCGGCTCGT CAAGTTCGTC TTCCGGGAGC CGGAGGACGT CTGGCGCAGG GCAGTCCGTC TCCAATAGCG GGGACGACAG

+2 Q T N W Q K L E T F W A K H M W N F I S G I Q Y L A G
 3521 CAGACCAACT GGCAAAAACCT CGAGACCTTC TGGGCGAAGC ATATGTGGAA CTTTCATCAGT GGGATACAAT ACTTGGCGGG
 GTCTGGTTGA CCGTTTTTGA GCTCTGGAAG ACCCGCTTCG TATACACCTT GAAGTAGTCA CCCTATGTTA TGAACCGCCC

+2 L S T L P G N P A I A S L M A F T A A V T S P L T T
 3601 CTTGTCAACG CTGCTGGTA ACCCGCCAT TGCTTCATTG ATGGCTTTTA CAGCTGCTGT CACCAGCCCA CTAACCACTA
 GAACAGTTGC GACGGACCAT TGGGGCGGTA ACGAAGTAAC TACCGAAAAT GTCGACGACA GTGGTGGGT .GATTGGTGT

+2 S Q T L L F N I L G G W V A A Q L A A P G A A T A F V
 3681 GCCAAACCCT CCTCTTCAAC ATATTGGGGG GGTGGGTGGC TGCCAGCTC GCCGCCCCCG GTGCGGCTAC TGCCTTTGTG
 CGGTTTGGGA GGAGAAGTTG TATAACCCCC CCACCCACCG ACGGTGTCAG CGGCGGGGGC CACGGCGATG ACGGAAACAC

FIGURE 5 - Page 4

+2 G A G L A G A A I G S V G L G K V L I D I L A G Y G A
 3761. GGCGCTGGCT TAGCTGGCGC CGCCATCGGC AGTGTGGAC TGGGGAAGGT CCTCATAGAC ATCCTTGAG GGTATGGCGC
 CCGCGACCGA ATCGACCGC GCGGTAGCCG TCACAACCTG ACCCCTTCCA GGAGTATCTG TAGGAACGTC CCATACCGCG

+2 G V A G A L V A F K I M S G E V P S T E D L V N L L
 3841. GGGCGTGGCG GGAGCTCTTG TGGCATTCAA GATCATGAGC GGTGAGGTCC CCTCCACGGA GGACCTGGTC AATCTACTGC
 CCCGCACCGC CCTCGAGAAC ACCGTAAGTT CTAGTACTCG CCACTCCAGG GGAGGTGCCT CCTGGACCAG TTAGATGACG

+2 P A I L S P G A L V V G V V C A A I L R R H V G P G E
 3921. CCGCCATCCT CTCGCCCCGGA GCCCTCGTAG TCGGCGTGGT CTGTGCAGCA ATACTGCGCC GGCACGTTGG CCCGGGCGAG
 GGGCGTAGGA GAGCGGGCCT CGGGAGCATC AGCCGCACCA GACACGTCGT TATGACGCGG CCGTGCAACC GGGCCCGCTC

+2 G A V Q W M N R L I A F A S R G N H V S P T H Y V P E
 4001. GGGGCGAGTGC AGTGGATGAA CCGGCTGATA GCCTTCGCCT CCCGGGGGAA CCATGTTTCC CCCACGCACT ACGTGCCGGA
 CCCCCTCAGC TCACCTACTT GGCCGACTAT CGGAAGCGGA GGGCCCCCTT GGTACAAAGG GGGTGCGTGA TGCACGGCCT

+2 S D A A A R V T A I L S S L T V T Q L L R R L H Q W
 4081. GAGCGATGCA GCTGCCCCGCG TCACTGCCAT ACTCAGCAGC CTCACTGTAA CCCAGCTCCT GAGGCGACTG CACCACTGGA
 CTCGTACGT CGACGGGCGC AGTGACGGTA TGAGTCGTCG GAGTGACATT GGGTCGAGGA CTCCGTGAC GTGGTCACCT

+2 I S S E C T T P C S G S W L R D I W D W I C E V L S D
 4161. TAAGCTCGGA GTGTACCACT CCATGCTCCG GTTCTGGCT AAGGGACATC TGGGACTGGA TATGCGAGGT GTTGAGCGAC
 ATTCGAGCCT CACATGGTGA GGTACGAGGC CAAGGACCGA TTCCCTGTAG ACCCTGACCT ATACGCTCCA CAACTCGCTG

+2 F K T W L K A K L M P Q L P G I P F V S C Q R G Y K G
 BamHI
 ~~~~~

4241. TTTAAGACCT GGCTAAAAGC TAAGCTCATG CCACAGCTGC CTGGGATCCC CTTTGTGTCC TGCCAGCGCG GGTATAAGGG  
 AAATTCTGGA CCGATTTTCG ATTCGAGTAC GGTGTGCGAC GACCCTAGGG GAAACACAGG ACGGTCGCGC CCATATTCCC

+2 V W R G D G I M H T R C H C G A E I T G H V K N G T  
 4321. GGTCTGGCGA GGGGACGGCA TCATGCACAC TCGCTGCCAC TGTGGAGCTG AGATCACTGG ACATGTCAAA AACGGGACGA  
 CCAGACCGCT CCCCTGCCGT AGTACGTGTG AGCGACGGTG ACACCTCGAC TCTAGTGACC TGTACAGTTT TTGCCCTGCT

+2 M R I V G P R T C R N M W S G T F P I N A Y T T G P C  
 4401. TGAGGATCGT CGGTCTTAGG ACCTGCAGGA ACATGTGGAG TGGGACCTTC CCCATTAATG CCTACACCAC GGGCCCCGTG  
 ACTCTAGCA GCCAGGATCC TGGACGTCCT TGTACACCTC ACCCTGGAAG GGGTAATTAC GGATGTGGTG CCCGGGGACA

+2 T P L P A P N Y T F A L W R V S A E E Y V E I R Q V G  
 4481. ACCCCCCCTT CTGCGCCGAA CTACACGTTT GCGCTATGGA GGGTGTCTGC AGAGGAATAC GTGGAGATAA GGCAGGTGGG  
 TGGGGGGAAG GACGCGGCTT GATGTGCAAG CGCGATACCT CCCACAGACG TCTCCTTATG CACCTCTATT CCGTCCACCC

+2 D F H Y V T G M T T D N L K C P C Q V P S P E F F T  
 4561. GGAATTCAC TACGTGACGG GTATGACTAC TGACAATCTT AAATGCCCCT GCCAGTCCC ATCGCCCGAA TTTTTCACAG  
 CCTGAAGGTG ATGCACTGCC CATACTGATG ACTGTTAGAA TTTACGGGCA CGGTCCAGGG TAGCGGGCTT AAAAAGTGTC

+2 E L D G V R L H R F A P P C K P L L R E E V S F R V G  
 4641. AATTGGACGG GGTGCGCCTA CATAGGTTTG CGCCCCCTG CAAGCCCTTG CTGCGGGAGG AGGTATCATT CAGAGTAGGA  
 TTAACCTGCC CCACGCGGAT GTATCCAAAC GCGGGGGGAC GTTCGGGAAC GACGCCCTCC TCCATAGTAA GTCTCATCCT

+2 L H E Y P V G S Q L P C E P E P D V A V L T S M L T D  
 4721. CTCCACGAAT ACCCGGTAGG GTCGCAATTA CCTTGCGAGC CCGAACCAGG CGTGCCCGTG TTGACGTCCA TGCTCACTGA  
 GAGGTGCTTA TGGGCCATCC CAGCGTTAAT GGAACGCTCG GGCTTGGCCT GCACCGGCAC AACTGCAGGT ACGAGTGACT

+2 P S H I T A E A A G R R L A R G S P P S V A S S S A  
 4801. TCCCTCCCAT ATAACAGCAG AGGCGGGCCG GCGAAGGTTG GCGAGGGGAT CACCCCCCTC TGTGGCCAGC TCCTCGGCTA  
 AGGGAGGGTA TATTGTCGTC TCCGCCGGCC CGCTTCCAAC CGCTCCCCTA GTGGGGGGAG ACACCGGTG AGGAGCCGAT

5921 +2 V T P I D T T I M A K N E V F C V Q P E K G G R K P A  
GTAACACCAA TAGACACTAC CATCATGGCT AAGAACGAGG TTTTCTGCGT TCAGCCTGAG AAGGGGGGTC GTAAGCCAGG  
CATTGTGGTT ATCTGTGATG GTAGTACCGA TTCTTGCTCC AAAAGACGCA AGTCGGACTC TTCCCCCAG CATTGGGTCC

[illegible]



## FIGURE 5 - Page 6

+2 R L I V F P D L G V R V C E K M A L Y D V V T K L P  
 6001 TCGTCTCATC GTGTTCCCGG ATCTGGGCGT GCGCGTGTGC GAAAAGATGG CTTTGTACGA CGTGGTTACA AAGCTCCCCT  
 AGCAGAGTAG CACAAGGGGC TAGACCCGCA CGCGCACACG CTTTCTACC GAAACATGCT GCACCAATGT TTCGAGGGGA

---

+2 L A V M G S S Y G F Q Y S P G Q R V E F L V Q A W K S  
 EcoRI  
 ~~~~~  
 6081 TGGCCGTGAT GGAAGCTCC TACGGATTCC AATACTCACC AGGACAGCGG GTTGAATTCC TCGTGCAAGC GTGGAAGTCC
 ACCGGCACTA CCCTTCGAGG ATGCCTAAGG TTATGAGTGG TCCTGTCGCC CAACTTAAGG AGCACGTTCC CACCTTCAGG

+2 K K T P M G F S Y D T R C F D S T V T E S D I R T E E
 6161 AAGAAAACCC CAATGGGGTT CTCGTATGAT ACCCGCTGCT TTGACTCCAC AGTCACTGAG AGCGACATCC GTACGGAGGA
 TTCTTTTGGG GTTACCCCAA GAGCATACTA TGGGCGACGA AACTGAGGTG TCAGTGACTC TCGCTGTAGG CATGCCTCCT

+2 A I Y Q C C D L D P Q A R V A I K S L T E R L Y V G
 6241 GGCAATCTAC CAATGTTGTG ACCTCGACCC CCAAGCCCGC GTGGCCATCA AGTCCCTCAC CGAGAGGCTT TATGTTGGGG
 CCGTTAGATG GTTACAACAC TGGAGCTGGG GGTTCCGGGCG CACCGGTAGT TCAGGGAGTG GCTCTCCGAA ATACAACCC

+2 G P L T N S R G E N C G Y R R C R A S G V L T T S C G
 6321 GCCCTCTTAC CAATTCAAGG GGGGAGAACT GCGGCTATCG CAGGTGCCGC GCGAGCGGCG TACTGACAAC TAGCTGTGGT
 CGGGAGAATG GTTAAGTTCC CCCCTCTTGA CGCGATAGC GTCCACGGCG CGCTCGCCGC ATGACTGTTG ATCGACACCA

+2 N T L T C Y I K A R A A C R A A G L Q D C T M L V C G
 6401 AACACCCTCA CTTGCTACAT CAAGGCCCGG GCAGCCTGTC GAGCCGAGG GCTCCAGGAC TGCACCATGC TCGTGTGTGG
 TTGTGGGAGT GAACGATGTA GTTCCGGGCC CGTCGGACAG CTCGGCGTCC CGAGGTCCTG ACGTGGTACG AGCACACACC

+2 D D L V V I C E S A G V Q E D A A S L R A F T E A M
 6481 CGACGACTTA GTCGTTATCT GTGAAAGCGC GGGGGTCCAG GAGGACGCGG CGAGCCTGAG AGCCTTCACG GAGGCTATGA
 GCTGCTGAAT CAGCAATAGA CACTTTCGCG CCCCAGGTC CTCCTGCGCC GCTCGGACTC TCGGAAGTGC CTCCGATACT

+2 T R Y S A P P G D P P Q P E Y D L E L I T S C S S N V
 6561 CCAGGTACTC CGCCCCCCT GGGGACCCCG CACAACCAGA ATACGACTTG GAGCTCATAA CATCATGCTC CTCCAACGTG
 GGTCCATGAG GCGGGGGGGA CCCCTGGGGG GTGTTGGTCT TATGCTGAAC CTCGAGTATT GTAGTACGAG GAGGTTGCAC

+2 S V A H D G A G K R V Y Y L T R D P T T P L A R A A W
 6641 TCAGTCGCCC ACGACGGCGC TGGAAAGAGG GTCTACTACC TCACCCGTGA CCCTACAACC CCCCTCGCGA GAGCTGCGTG
 AGTCAGCGGG TGCTGCCGCG ACCTTTCTCC CAGATGATGG AGTGGGCACT GGGATGTTGG GGGGAGCGCT CTCGACGCAC

+2 E T A R H T P V N S W L G N I I M F A P T L W A R M
 6721 GGAGACAGCA AGACACACTC CAGTCAATTC CTGGCTAGGC AACATAATCA TGTTTGCCCC CACACTGTGG GCGAGGATGA
 CCTCTGTCGT TCTGTGTGAG GTCAGTTAAG GACCGATCCG TTGTATTAGT ACAAACGGGG GTGTGACACC CGCTCCTACT

+2 I L M T H F F S V L I A R D Q L E Q A L D C E I Y G A
 6801 TACTGATGAC CCATTTCTTT AGCGTCCTTA TAGCCAGGGA CCAGCTTGAA CAGGCCCTCG ATTGCGAGAT CTACGGGGCC
 ATGACTACTG GGTAAAGAAA TCGCAGGAAT ATCGGTCCTT GGTCGAACTT GTCCGGGAGC TAACGCTCTA GATGCCCCGG

+2 C Y S I E P L D L P P I I Q R L H G L S A F S L H S Y
 6881 TGCTACTCCA TAGAACCACT GGATCTACCT CCAATCATTC AAAGACTCCA TGGCCTCAGC GCATTTTCAC TCCACAGTTA
 ACGATGAGGT ATCTTGGTGA CCTAGATGGA GGTTAGTAAG TTTCTGAGGT ACCGGAGTCG CGTAAAAGTG AGGTGTCAAT

+2 S P G E I N R V A A C L R K L G V P P L R A W R H R
 6961 CTCTCCAGGT GAAATCAATA GGGTGGCCGC ATGCCTCAGA AAACCTGGGG TACCGCCCTT GCGAGCTTGG AGACACGGGG
 GAGAGGTCCA CTTTAGTTAT CCCACGGGCG TACGGAGTCT TTTGAACCCC ATGGCGGGAA CGCTCGAACC TCTGTGGCCC

+2 A R S V R A R L L A R G G R A A I C G K Y L F N W A V
 7041 CCCGGAGCGT CCGCGCTAGG CTTCTGGCCA GAGGAGGCAG GGCTGCCATA TGTGGCAAGT ACCTCTTCAA CTGGGCAGTA
 GGGCCTCGCA GCGCGATCC GAAGACCGGT CTCCTCCGTC CCGACGGTAT ACACCGTTCA TGGAGAAGTT GACCCGTCAT

8401 TGATCTTTTC TACGGGTCT GACGCTCAGT GGAACGAAAA CTCACGTTAA GGGATTTTGG TCATGAGATT ATCAAAAAGG
ACTAGAAAAG ATGCCCCAGA CTGCGAGTCA CCTTGCTTTT GAGTGCAATT CCCTAAAACC AGTACTCTAA TAGTTTTTCC

FIGURE 5 - Page 8

8481 ATCTTCACCT AGATCCTTTT AAATTAAAAA TGAAGTTTTA AATCAATCTA AAGTATATAT GAGTAAACTT GGTCTGACAG
TAGAAGTGGA TCTAGGAAAA TTTAATTTTT ACTTCAAAAT TTAGTTAGAT TTCATATATA CTCATTTGAA CCAGACTGTC

8561 TTACCAATGC TTAATCAGTG AGGCACCTAT CTCAGCGATC TGTCTATTTT GTTCATCCAT AGTTGCCTGA CTCCCCGTCTG
AATGGTTACG AATTAGTCAC TCCGTGGATA GAGTCGCTAG ACAGATAAAG CAAGTAGGTA TCAACGGACT GAGGGGCAGC

8641 TGTAGATAAC TACGATACGG GAGGGCTTAC CATCTGGCCC CAGTGCTGCA ATGATACCGC GAGACCCACG CTCACCGGCT
ACATCTATTG ATGCTATGCC CTCCCGAATG GTAGACCGGG GTCACGACGT TACTATGGCG CTCTGGGTGC GAGTGGCCGA

8721 CCAGATTTAT CAGCAATAAA CCAGCCAGCC GGAAGGGCCG AGCGCAGAAG TGGTCCTGCA ACTTTATCCG CCTCCATCCA
GGTCTAAATA GTCGTTATTT GGTGCGTCGG CCTTCCCGGC TCGCGTCTTC ACCAGGACGT TGAAATAGGC GGAGGTAGGT

8801 GTCTATTAAT TGTGCGGGG AAGCTAGAGT AAGTAGTTCG CCAGTTAATA GTTTGCGCAA CGTTGTTGCC ATTGCTACAG
CAGATAATTA ACAACGGCCC TTCGATCTCA TTCATCAAGC GGTCAATTAT CAAACGCGTT GCAACAACGG TAACGATGTC

8881 GCATCGTGGT GTCACGCTCG TCGTTTGGTA TGGCTTCATT CAGCTCCGGT TCCCAACGAT CAAGGCGAGT TACATGATCC
CGTAGCACCA CAGTGCGAGC AGCAAACCAT ACCGAAGTAA GTCGAGGCCA AGGGTTGCTA GTTCCGCTCA ATGTACTAGG

8961 CCCATGTTGT GCAAAAAAGC GGTAGCTCC TTCGGTCCTC CGATCGTTGT CAGAAGTAAG TTGGCCGCGAG TGTATCACT
GGGTACAACA CGTTTTTTCG CCAATCGAGG AAGCCAGGAG GCTAGCAACA GTCTTCATT CAAACGCGTC ACAATAGTGA

9041 CATGGTTATG GCAGCACTGC ATAATTCTCT TACTGTCATG CCATCCGTAA GATGCTTTTC TGTGACTGGT GAGTACTCAA
GTACCAATAC CGTCGTGACG TATTAAGAGA ATGACAGTAC GGTAGGCATT CTACGAAAAG AACTGACCA CTCATGAGTT

9121 CCAAGTCATT CTGAGAATAG TGTATGCGGC GACCGAGTTG CTCTTGCCCG GCGTCAATAC GGGATAATAC CGCGCCACAT
GGTTCAGTAA GACTCTTATC ACATACGCCG CTGGCTCAAC GAGAACGGGC CGCAGTTATG CCCTATTATG GCGCGGTGTA

9201 AGCAGAACTT TAAAAGTGCT CATCATTGGA AAACGTTCTT CGGGGCGAAA ACTCTCAAGG ATCTTACCGC TGTGAGATC
TCGTCTTGAA ATTTTCACGA GTAGTAACCT TTTGCAAGAA GCCCCGCTTT TGAGAGTTCC TAGAATGGCG ACAACTCTAG

9281 CAGTTCGATG TAACCCACTC GTGCACCCAA CTGATCTTCA GCATCTTTTA CTTTCACCAG CGTTTCTGGG TGAGCAAAAA
GTCAAGCTAC ATTGGGTGAG CACGTGGGT GACTAGAAGT CGTAGAAAAT GAAAGTGGTC GCAAAGACCC ACTCGTTTTT

9361 CAGGAAGGCA AAATGCCGCA AAAAAGGGAA TAAGGGCGAC ACGGAAATGT TGAATACTCA TACTCTTCCT TTTTCAATAT
GTCCTTCCGT TTTACGGCGT TTTTCCCTT ATTCCCGCTG TGCCCTTACA ACTTATGAGT ATGAGAAGGA AAAAGTTATA

9441 TATTGAAGCA TTTATCAGGG TTATTGTCTC ATGAGCGGAT ACATATTTGA ATGTATTTAG AAAAATAAAC AAATAGGGGT
ATAACTTCGT AAATAGTCCC AATAACAGAG TACTCGCCTA TGTATAAAT TACATAAATC TTTTATTTG TTTATCCCCA

9521 TCCGCGCACA TTTCCCCGAA AAGTGCCACC TGACGTCTAA GAAACCATTA TTATCATGAC ATTAACCTAT AAAAATAGGC
AGGCGCGTGT AAAGGGGCTT TTCACGGTGG ACTGCAGATT CTTTGTAAT AATAGTACTG TAATTGGATA TTTTATCCG

9601 GTATCACGAG GCCCTTTCGT C
CATAGTGCTC CGGAAAAGCA G

09724764250

FIGURE 6

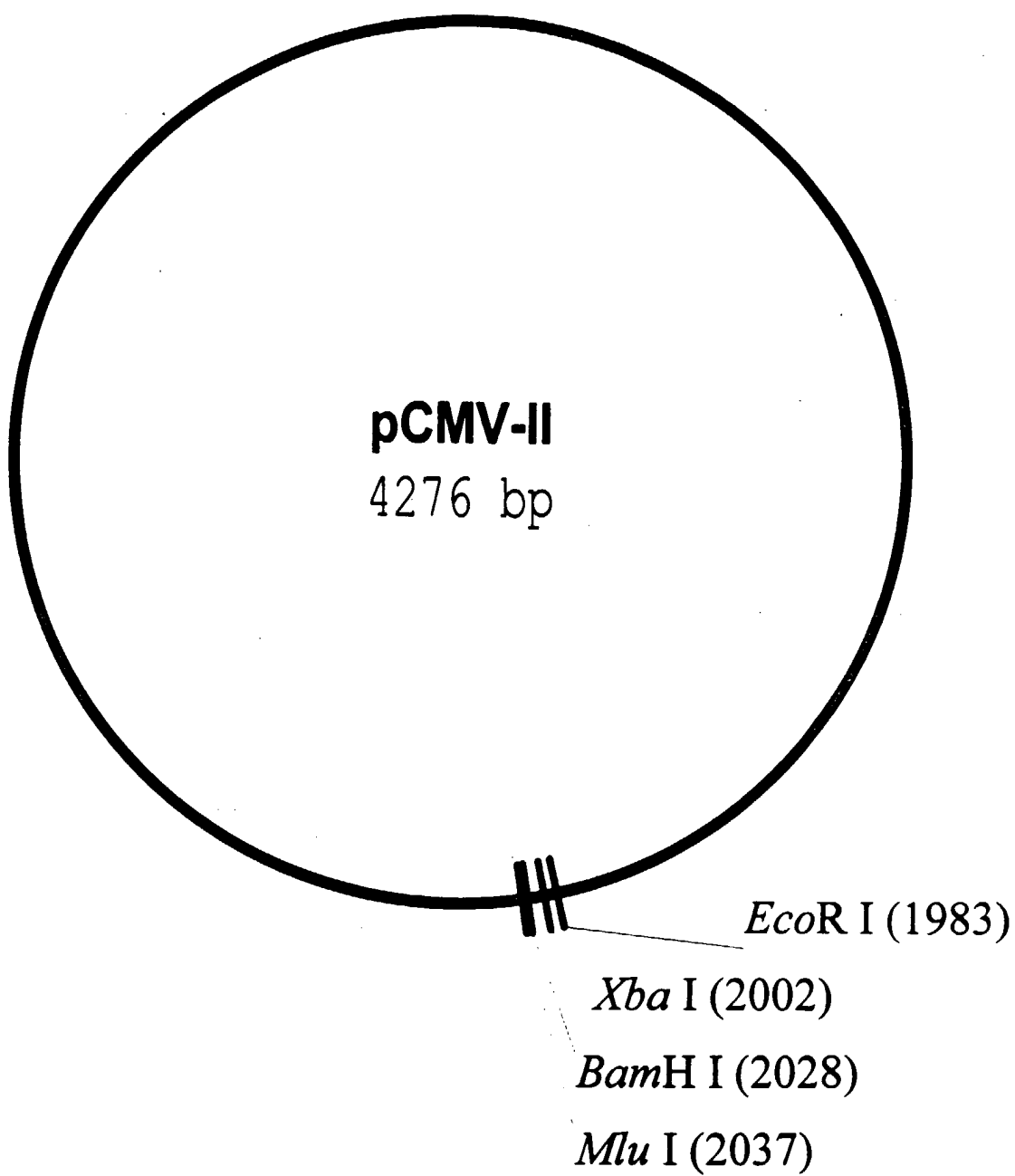


FIGURE 7 - Page 1

1	TCGCGCGTTT	CGGTGATGAC	GGTGAAGAAC	TCTGACACAT	GCAGCTCCCG	GAGACGGTCA	CAGCTTGTCT	GTAAGCGGAT
	AGCGCGCAAA	GCCACTACTG	CCACTTTTGG	AGACTGTGTA	CGTCGAGGGC	CTCTGCCAGT	GTCGAACAGA	CATTGCGCTA
81	GCCGGGAGCA	GACAAGCCCC	TCAGGGCGCG	TCAGCGGGTG	TTGGCGGGTG	TCGGGGCTGG	CTTAACATATG	CGGCATCAGA
	CGGCCCTCGT	CTGTTCCGGC	AGTCCCGCGC	AGTCGCCAC	AACCGCCCAC	AGCCCCGACC	GAATTGATAC	GCCGTAGTCT
161	GCAGATTGTA	CTGAGAGTGC	ACCATATGAA	GCTTTTTGCA	AAAGCCTAGG	CCTCCAAAAA	AGCCTCCTCA	CTACTTCTGG
	CGTCTAACAT	GACTCTCACG	TGGTATACTT	CGAAAAACGT	TTTCGGATCC	GGAGGTTTTT	TCGGAGGAGT	GATGAAGACC
241	AATAGCTCAG	AGGCCGAGGC	GGCCTCGGCC	TCTGCATAAA	TAAAAAAAAT	TAGTCAGCCA	TGGGGCGGAG	AATGGGCGGA
	TTATCGAGTC	TCCGGCTCCG	CCGGAGCCGG	AGACGTATTT	ATTTTTTTTA	ATCAGTCGGT	ACCCCGCCTC	TTACCCGCCT
321	ACTGGGCGGG	GAGGGAATTA	TTGGCTATTG	GCCATTGCAT	ACGTTGTATC	TATATCATAA	TATGTACATT	TATATTGGCT
	TGACCCGCCC	CTCCCTTAAT	AACCGATAAC	CGGTAACGTA	TGCAACATAG	ATATAGTATT	ATACATGTAA	ATATAACCGA
401	CATGTCCAAT	ATGACCGCCA	TGTTGACATT	GATTATTGAC	TAGTTATTAA	TAGTAATCAA	TTACGGGGTC	ATTAGTTCAT
	GTACAGGTTA	TACTGGCGGT	ACAACGTAA	CTAATAACTG	ATCAATAATT	ATCATTAGTT	AATGCCCCAG	TAATCAAGTA
481	AGCCCATATA	TGGAGTTCCG	CGTTACATAA	CTTACGGTAA	ATGGCCCGCC	TGGCTGACCG	CCCAACGACC	CCCGCCCAT
	TCGGGTATAT	ACCTCAAGGC	GCAATGTATT	GAATGCCATT	TACCGGGCGG	ACCGACTGGC	GGGTGCTGG	GGGCGGGTAA
561	GACGTCAATA	ATGACGTATG	TTCCCATAGT	AACGCCAATA	GGGACTTTCC	ATTGACGTCA	ATGGGTGGAG	TATTTACGGT
	CTGCAGTTAT	TACTGCATAC	AAGGGTATCA	TTGCGGTTAT	CCCTGAAAGG	TAATGTCAGT	TACCCACCTC	ATAAATGCCA
641	AAACTGCCCC	CTTGGCAGTA	CATCAAGTGT	ATCATATGCC	AAGTCCGCCC	CCTATTGACG	TCAATGACGG	TAAATGGCCC
	TTTGACGGGT	GAACCGTCAT	GTAGTTCACA	TAGTATACGG	TTCAGGCGGG	GGATAACTGC	AGTTACTGCC	ATTTACCGGG
721	GCCTGGCATT	ATGCCAGTA	CATGACCTTA	CGGGACTTTC	CTACTTGGCA	GTACATCTAC	GTATTAGTCA	TCGCTATTAC
	CGGACCGTAA	TACGGGTCAT	GTAAGTGAAT	GCCCTGAAAG	GATGAACCGT	CATGTAGATG	CATAATCAGT	AGCGATAATG
801	CATGGTGATG	CGGTTTTTGG	AGTACACCAA	TGGGCGTGGA	TAGCGGTTTG	ACTCACGGGG	ATTTCCAAGT	CTCCACCCCA
	GTACCACTAC	GCCAAAACCG	TCATGTGGTT	ACCCGCACCT	ATCGCCAAAC	TGAGTGCCCC	TAAAGGTTCA	GAGGTGGGGT
881	TTGACGTCAA	TGGGAGTTTG	TTTTGGCACC	AAAATCAACG	GGACTTTCCA	AAATGTCGTA	ATAACCCCGC	CCCGTTGACG
	AACGTGAGTT	ACCCTCAAAC	AAAACCGTGG	TTTAGTTGTC	CCTGAAAGGT	TTTACAGCAT	TATTGGGGCG	GGGCAACTGC
961	CAAATGGGCG	GTAGGCGTGT	ACGGTGGGAG	GTCTATATAA	GCAGAGCTCG	TTTAGTGAAC	CGTCAGATCG	CCTGGAGACG
	GTTTACCCGC	CATCCGCACA	TGCCACCCTC	CAGATATATT	CGTCTCGAGC	AAATCACTTG	GCAGTCTAGC	GGACCTCTGC
1041	CCATCCACGC	TGTTTTGACC	TCCATAGAAG	ACACCGGGAC	CGATCCAGCC	TCCGCGGCGG	GGAACGGTGC	ATTGGAACGC
	GGTAGGTGCG	ACAAAACCTG	AGGTATCTTC	TGTGGCCCTG	GCTAGGTCCG	AGGCGCCGGC	CCTTGCCACG	TAACCTTGCG
1121	GGATTCCCCG	TGCCAAGAGT	GACGTAAGTA	CCGCTATAG	ACTCTATAGG	CACACCCCTT	TGGCTCTTAT	GCATGCTATA
	CCTAAGGGGC	ACGGTTCTCA	CTGCATTCTA	GGCGGATATC	TGAGATATCC	GTGTGGGGAA	ACCGAGAATA	CGTACGATAT
1201	CTGTTTTTGG	CTTGGGGCCT	ATACACCCCC	GCTCCTTATG	CTATAGGTGA	TGGTATAGCT	TAGCCTATAG	GTGTGGGGTA
	GACAAAACCC	GAACCCCGGA	TATGTGGGGG	CGAGGAATAC	GATATCCACT	ACCATATCGA	ATCGGATATC	CACACCCAAT
1281	TTGACCATTA	TTGACCACTC	CCCTATTGGT	GACGATACTT	TCCATTACTA	ATCCATAACA	TGGCTCTTTG	CCACAACAT
	AACGTGTAAT	AACGTGTGAG	GGGATAACCA	CTGCTATGAA	AGGTAATGAT	TAGGTATTGT	ACCGAGAAAC	GGTGTGATA
1361	CTCTATTGGC	TATATGCCAA	TACTCTGTCC	TTCAGAGACT	GACACGGACT	CTGTATTTTT	ACAGGATGGG	GTCCATTAT
	GAGATAACCG	ATATACGGTT	ATGAGACAGG	AAGTCTCTGA	CTGTGCCTGA	GACATAAAAA	TGTCTACCC	CAGGTAAATA
1441	TATTTACAAA	TTCACATATA	CAACAACGCC	GTCCCCCGTG	CCCGCAGTTT	TTATTAAACA	TAGCGTGGGA	TCTCCGACAT
	ATAAATGTTT	AAGTGTATAT	GTTGTTGCGG	CAGGGGGCAC	GGGCGTCAAA	AATAATTTGT	ATCGCACCT	AGAGGCTGTA

1521	CTCGGGTACG GAGCCCATGC	TGTTCCGGAC ACAAGGCCTG	ATGGGCTCTT TACCCGAGAA	CTCCGGTAGC GAGGCCATCG	GGCGGAGCTT CCGCCTCGAA	CCACATCCGA GGTGTAGGCT	GCCCTGGTCC CGGGACCAGG	CATCCGTCCA GTAGGCAGGT
1601	GCGGCTCATG CGCCGAGTAC	GTCGCTCGGC CAGCGAGCCG	AGCTCCTTGC TCGAGGAACG	TCCTAACAGT AGGATTGTCA	GGAGGCCAGA CCTCCGGTCT	CTTAGGCACA GAATCCGTGT	GCACAAATGCC CGTGTCTACG	CACCACCACC GTGGTGGTGG
1681	AGTGTGCCGC TCACACGGCG	ACAAGGCCGT TGTTCCGGCA	GGCGGTAGGG CCGCCATCCC	TATGTGTCTG ATACACAGAC	AAAATGAGCT TTTACTCGA	CGGAGATTGG GCCTCTAACC	GCTCGCACCT CGAGCGTGA	GGACGCAGAT CCTGCGTCTA
1761	GGAAGACTTA CCTTCTGAAT	AGGCAGCGGC TCCGTCGCCG	AGAAGAAGAT TCTTCTTCTA	GCAGGCAGCT CGTCCGTCGA	GAGTTGTTGT CTCAACAACA	ATTCTGATAA TAAGACTATT	GAGTCAGAGG CTCAGTCTCC	TAACCTCCCGT ATTGAGGGCA
1841	TGCGGTGCTG ACGCCACGAC	TTAACGGTGG AATTGCCACC	AGGGCAGTGT TCCCGTCACA	AGTCTGAGCA TCAGACTCGT	GTAATCGTTG CATGAGCAAC	CTGCCGCGCG GACGGCGCGC	CGCCACCAGA GCGGTGGTCT	CATAATAGCT GTATTATCGA
1921	GACAGACTAA CTGTCTGATT	CAGACTGTTT GTCTGACAAG	CTTTCCATGG GAAAGGTACC	GTCTTTTCTG CAGAAAAGAC	CAGTCACCGT GTCAGTGCCA	CGTCGACCTA GCAGCTGGAT	EcoRI ~~~~~ AGAATTGAGA TCTTAAGTCT	CTCGAGCAAG GAGCTCGTTC
2001	XbaI ~~~~~ TCTAGAAAGG AGATCTTTCC	CGCGCCAAGA GCGCGGTTCT	BamHI ~~~~~ TATCAAGGAT ATAGTTCTTA	MluI ~~~~~ CCACTACGCG GGTGATGCGC	TTAGAGCTCG AATCTCGAGC	CTGATCAGCC GACTAGTCGG	TCGACTGTGC AGCTGACACG	CTTCTAGTTG GAAGATCAAC
2081	CCAGCCATCT GGTCGGTAGA	GTTGTTTGCC CAACAAACGG	CCTCCCCCGT GGAGGGGGCA	GCCTTCCTTG CGGAAGGAAC	ACCCTGGAAG TGGGACCTTC	GTGCCACTCC CACGGTGAGG	CACTGTCCTT GTGACAGGAA	TCCTAATAAA AGGATTATTT
2161	ATGAGGAAAT TACTCCTTTA	TGCATCGCAT ACGTAGCGTA	TGTCTGAGTA ACAGACTCAT	GGTGTCAATC CCACAGTAAG	TATTCTGGGG ATAAGACCCC	GGTGGGGTGG CCACCCCACC	GGCAGGACAG CCGTCTGTGC	CAAGGGGGAG GTTCCCCCTC
2241	GATTGGGAAG CTAACCCCTT	ACAATAGCAG TGTTATCGTC	GCATGCTGGG CGTACGACCC	GAGCTCTTCC CTCGAGAAAG	GCTTCCTCGC CGAAGGAGCG	TCACTGACTC AGTGACTGAG	GCTGCGCTCG CGACGCGAGC	GTCGTTCCGG CAGCAAGCCG
2321	TGCGGCGAGC ACGCCGCTCG	GGTATCAGCT CCATAGTCGA	CACTCAAAGG GTGAGTTTCC	CGGTAATACG GCCATTATGC	GTTATCCACA CAATAGGTGT	GAATCAGGGG CTTAGTCCCC	ATAACGCAGG TATTGCGTCC	AAAGAACATG TTCTTGTAC
2401	TGAGCAAAAAG ACTCGTTTTT	GCCAGCAAAA CGGTCGTTTT	GGCCAGGAAC CCGGTCCTTG	CGTAAAAAGG GCATTTTTTCC	CCGCGTTGCT GGCGCAACGA	GGCGTTTTTC CCGCAAAAAG	CATAGGCTCC GTATCCGAGG	GCCCCCTGA CGGGGGGACT
2481	CGAGCATCAC GCTCGTAGTG	AAAAATCGAC TTTTTAGCTG	GCTCAAGTCA CGAGTTCAGT	GAGGTGGCGA CTCCACCGCT	AACCCGACAG TTGGGCTGTC	GACTATAAAG CTGATATTTT	ATACCAGGCG TATGGTCCGC	TTTCCCCCTG AAAGGGGGAC
2561	GAAGCTCCCT CTTCGAGGGA	CGTGCCTCT GCACGCGAGA	CCTGTTCCGA GGACAAGGCT	CCCTGCCGCT GGGACGGCGA	TACCGGATAC ATGGCCTATG	CTGTCCGCT GACAGGCGGA	TTCTCCCTTC AAGAGGGAAG	GGGAAGCGTG CCCTTCGCAC
2641	GCGCTTTCTC CGCGAAAGAG	AATGCTCACG TTACGAGTGC	CTGTAGGTAT GACATCCATA	CTCAGTTCGG GAGTCAAGCC	TGTAGGTCGT ACATCCAGCA	TCGCTCCAAG AGCGAGGTTC	CTGGGCTGTG GACCCGACAC	TGCACGAACC ACGTGCTTGG
2721	CCCCGTTTCT GGGGCAAGTC	CCCGACCGCT GGGCTGGCGA	GCGCCTTATC CGCGGAATAG	CGGTAACAT GCCATTGATA	CGTCTTGAGT GCAGAACTCA	CCAACCCGGT GGTTGGGCCA	AAGACACGAC TTCTGTGCTG	TTATCGCCAC AATAGCGGTG
2801	TGGCAGCAGC ACCGTCGTCG	CACTGGTAAC GTGACCATTG	AGGATTAGCA TCCTAATCGT	GAGCGAGGTA CTCGCTCCAT	TGTAGGCGGT ACATCCGCCA	GCTACAGAGT CGATGTCTCA	TCTTGAAGTG AGAACTTCAC	GTGGCCTAAC CACCGGATTG
2881	TACGGCTACA ATGCCGATGT	CTAGAAGGAC GATCTTCTCT	AGTATTTGGT TCATAAACCA	ATCTGCGCTC TAGACGCGAG	TGCTGAAGCC ACGACTTCGG	AGTTACCTTC TCAATGGAAG	GGAAAAAGAG CCTTTTTTCT	TTGGTAGCTC AACCATCGAG

FIGURE 7 - Page 3

2961 TTGATCCGGC AAACAAACCA CCGCTGGTAG CCGTGGTTTT TTTGTTTGCA AGCAGCAGAT TACGCGCAGA AAAAAAGGAT
AACTAGGCCG TTTGTTTGGT GCGGACCATC GCCACCAAAA AAACAAACGT TCGTCGTCTA ATGCGCGTCT TTTTTCCTA

3041 CTCAAGAAGA TCCTTTGATC TTTTCTACGG GGTCTGACGC TCAGTGGAAC GAAACTCAC GTTAAGGGAT TTGGTTCATG
GAGTTCTTCT AGGAAACTAG AAAAGATGCC CCAGACTGCG AGTCACCTTG CTTTGTAGTG CAATTCCCTA AAACCAGTAC

3121 AGATTATCAA AAAGGATCTT CACCTAGATC CTTTTAAATT AAAAATGAAG TTTTAAATCA ATCTAAAGTA TATATGAGTA
TCTAATAGTT TTTCTAGAA GTGGATCTAG GAAAATTTAA TTTTACTTC AAAATTTAGT TAGATTTTCAT ATATACTCAT

3201 AACTTGGTCT GACAGTTACC AATGCTTAAT CAGTGAGGCA CCTATCTCAG CGATCTGTCT ATTTTCGTTCA TCCATAGTTG
TTGAACCAGA CTGTCAATGG TTACGAATTA GTCACTCCGT GGATAGAGTC GCTAGACAGA TAAAGCAAGT AGGTATCAAC

3281 CCTGACTCCC CGTCGTGTAG ATAACACGA TACGGGAGGG CTTACCATCT GGCCCCAGTG CTGCAATGAT ACCGCGAGAC
GGACTGAGGG GCAGCACATC TATTGATGCT ATGCCCTCCC GAATGGTAGA CCGGGGTCAC GACGTTACTA TGGCGCTCTG

3361 CCACGCTCAC CGGCTCCAGA TTTATCAGCA ATAAACCAGC CAGCCGGAAG GGCCGAGCGC AGAAGTGGTC CTGCAACTTT
GGTGCGAGTG GCCGAGGTCT AAATAGTCGT TATTTGGTCG GTCGGCCTTC CCGGCTCGCG TCTTACCAG GACGTTGAAA

3441 ATCCGCCTCC ATCCAGTCTA TTAATTGTTG CCGGGAAGCT AGAGTAAGTA GTTCGCCAGT TAATAGTTTG CGCAACGTTG
TAGGCGGAGG TAGGTCAGAT AATTAACAAC GGCCCTTCGA TCTCATTCTA CAAGCGGTCA ATTATCAAAC GCGTTGCAAC

3521 TTGCCATTGC TACAGGCATC GTGGTGTAC GCTCGTCGTT TGGTATGGCT TCATTACAGT CCGGTTCCCA ACGATCAAGG
AACGGTAACG ATGTCGTAG CACCACAGTG CGAGCAGCAA ACCATACCGA AGTAAGTCGA GGCCAAGGGT TGCTAGTTCC

3601 CGAGTTACAT GATCCCCAT GTTGTGCAAA AAAGCGGTTA GCTCCTTCGG TCCTCCGATC GTTGTACAGAA GTAAGTTGGC
GCTCAATGTA CTAGGGGGTA CAACACGTTT TTTCCCAAT CGAGGAAGCC AGGAGGCTAG CAACAGTCTT CATTCAACCG

3681 CGCAGTGTTA TCACTCATGG TTATGGCAGC ACTGCATAAT TCTCTTACTG TCATGCCATC CGTAAGATGC TTTTCTGTGA
GCGTCACAAT AGTGAGTACC AATACCGTCG TGACGTATTA AGAGAATGAC AGTACGGTAG GCATTCTACG AAAAGACACT

3761 CTGGTGAGTA CTCAACCAAG TCATTCTGAG AATAGTGTAT GCGGCGACCG AGTTGCTCTT GCGGCGGTC AATACGGGAT
GACCACTCAT GAGTTGGTTC AGTAAGACTC TTATCACATA CGCCGCTGGC TCAACGAGAA CCGGCGCAG TTATGCCCTA

3841 AATACCGCGC CACATAGCAG AACTTTAAAA GTGCTCATCA TTGGAAAACG TTCTTCGGGG CGAAAACCTCT CAAGGATCTT
TTATGGCGCG GTGTATCGTC TTGAAATTTT CACGAGTAGT AACCTTTTGC AAGAAGCCCC GCTTTTGAGA GTTCCTAGAA

3921 ACCGCTGTTG AGATCCAGTT CGATGTAACC CACTCGTGCA CCCAACTGAT CTTACGATC TTTTACTTTC ACCAGCGTTT
TGGCGACAAC TCTAGGTCAA GCTACATTGG GTGAGCACGT GGGTTGACTA GAAGTCGTAG AAAATGAAAG TGGTCGCAAA

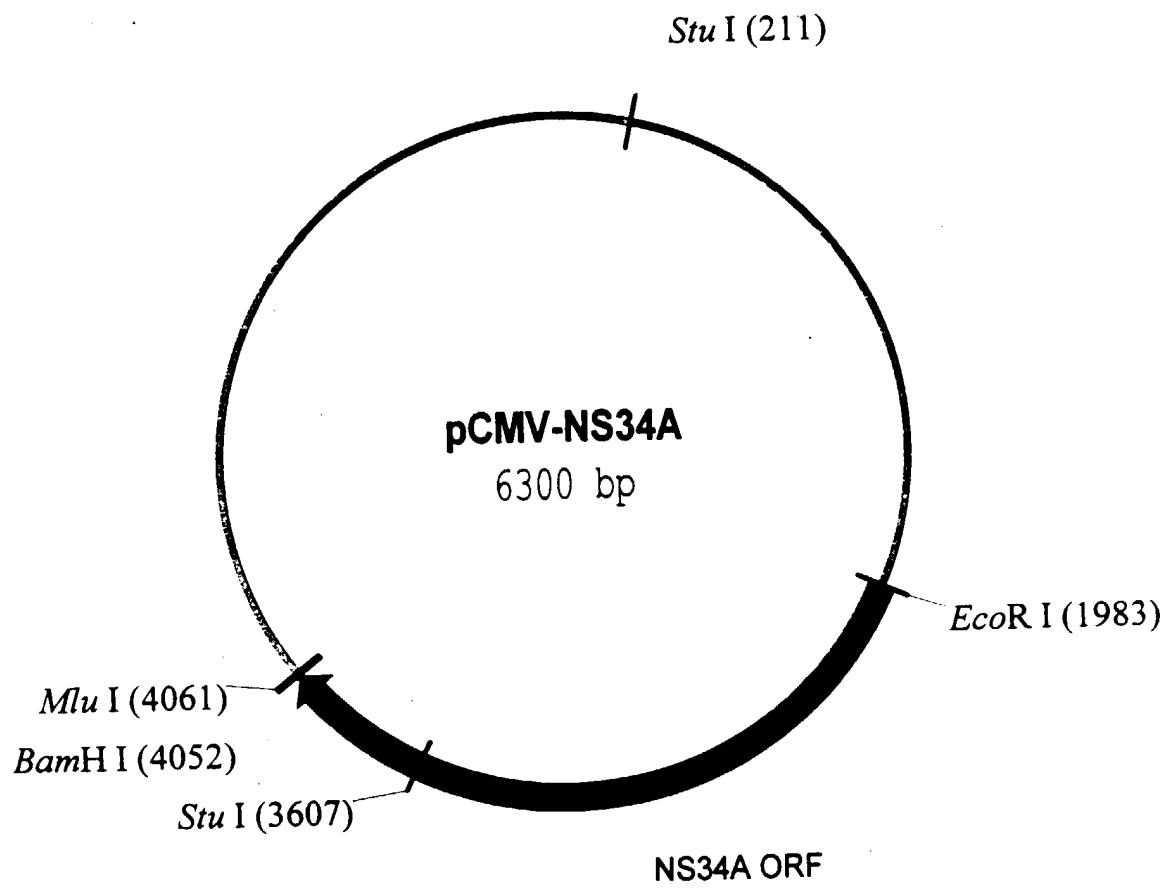
4001 CTGGGTGAGC AAAACAGGA AGGCAAAATG CCGCAAAAAA GGAATAAGG GCGACACGGA AATGTTGAAT ACTCATACTC
GACCACTCG TTTTGTCTCT TCCGTTTTAC GCGTTTTTTT CCCTTATTCC CGCTGTGCCT TTACAACCTA TGAGTATGAG

4081 TTCCTTTTTT AATATTATTG AAGCATTTAT CAGGGTTATT GTCTCATGAG CGGATACATA TTTGAATGTA TTTAGAAAAA
AAGGAAAAAG TTATAATAAC TTCGTAAATA GTCCCAATAA CAGAGTACTC GCCTATGTAT AAACCTTACAT AAATCTTTTT

4161 TAAACAAATA GGGGTTCCGC GCACATTTC CCGAAAAGTG CCACCTGACG TCTAAGAAAC CATTATTATC ATGACATTAA
ATTGTTTAT CCCCAAGGCG CGTGTAAGG GGCTTTTCAC GGTGGACTGC AGATTCTTTG GTAATAATAG TACTGTAATT

4241 CCTATAAAAA TAGGCGTATC ACGAGGCCCT TTCGTC
GGATATTTTT ATCCGCATAG TGCTCCGGGA AAGCAG

FIGURE 8



032376 6472650

FIGURE 9 - Page 1

1 TCGCGCGTTT CGGTGATGAC GGTGAAAACC TCTGACACAT GCAGCTCCCG
AGCGCGCAAA GCCACTACTG CCACTTTTGG AGACTGTGTA CGTCGAGGGC

51 GAGACGGTCA CAGCTTGTCT GTAAGCGGAT GCCGGGAGCA GACAAGCCCG
CTCTGCCAGT GTCGAACAGA CATTGCGCTA CGGCCCTCGT CTGTTGCGGG

101 TCAGGGCGCG TCAGCGGGTG TTGGCGGGTG TCGGGGCTGG CTTAACTATG
AGTCCCGCGC AGTCGCCCAC AACCGCCCAC AGCCCCGACC GAATTGATAC

151 CGGCATCAGA GCAGATTGTA CTGAGAGTGC ACCATATGAA GCTTTTGTGA
GCCGTAGTCT CGTCTAACAT GACTCTCAG TGGTATACTT CGAAAAACGT

StuI
~~~~~

201 AAAGCCTAGG CCTCCAAAAA AGCCTCCTCA CTACTTCTGG AATAGCTCAG  
TTTCGGATCC GGAGGTTTTT TCGGAGGAGT GATGAAGACC TTATCGAGTC

---

251 AGGCCGAGGC GGCTCGGCC TCTGCATAAA TAAAAAAAT TAGTCAGCCA  
TCCGGCTCCG CCGGAGCCGG AGACGTATTT ATTTTTTTTA ATCAGTCGGT

---

301 TGGGGCGGAG AATGGGCGGA ACTGGGCGGG GAGGGAATTA TTGGCTATTG  
ACCCCGCCTC TTACCCGCCT TGACCCGCCC CTCCCTTAAT AACCGATAAC

---

351 GCCATTGCAT ACGTTGTATC TATATCATAA TATGTACATT TATATTGGCT  
CGGTAACGTA TGCAACATAG ATATAGTATT ATACATGTAA ATATAACCGA

---

401 CATGTCCAAT ATGACCGCCA TGTTGACATT GATTATTGAC TAGTTATTAA  
GTACAGGTTA TACTGGCGGT ACAACTGTAA CTAATAACTG ATCAATAATT

---

451 TAGTAATCAA TTACGGGGTC ATTAGTTCAT AGCCCATATA TGGAGTTCCG  
ATCATTAGTT AATGCCCCAG TAATCAAGTA TCGGGTATAT ACCTCAAGGC

---

501 CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG CCCAACGACC  
GCAATGTATT GAATGCCATT TACCGGGCGG ACCGACTGGC GGGTTGCTGG

---

551 CCCGCCCAT TACGTCAATA ATGACGTATG TTCCCATAGT AACGCCAATA  
GGGCGGGTAA CTGCAGTTAT TACTGCATAC AAGGGTATCA TTGCGGTTAT

---

601 GGGACTTTCC ATTGACGTCA ATGGGTGGAG TATTTACGGT AACTGCCCCA  
CCCTGAAAGG TAACTGCAGT TACCCACCTC ATAAATGCCA TTTGACGGGT

---

651 CTTGGCAGTA CATCAAGTGT ATCATATGCC AAGTCCGCCC CCTATTGACG  
GAACCGTCAT GTAGTTCACA TAGTATACGG TTCAGGCGGG GGATAACTGC

---

701 TCAATGACGG TAAATGGCCC GCCTGGCATT ATGCCAGTA CATGACCTTA  
AGTTACTGCC ATTTACCGGG CGGACCGTAA TACGGGTCAT GTACTGGAAT

---

751 CGGGACTTTC CTACTTGGCA GTACATCTAC GTATTAGTCA TCGCTATTAC  
GCCCTGAAAG GATGAACCGT CATGTAGATG CATAATCAGT AGCGATAATG

---

801 CATGGTGATG CGGTTTTGGC AGTACACCAA TGGGCGTGGA TAGCGGTTTG  
GTACCACTAC GCCAAAACCG TCATGTGGTT ACCCGCACCT ATCGCCAAAC

---

851 ACTCACGGGG ATTTCCAAGT CTCCACCCCA TTGACGTCAA TGGGAGTTTG  
TGAGTGCCCC TAAAGGTTCA GAGGTGGGGT AACTGCAGTT ACCCTCAAAC

## FIGURE 9 - Page 2

901 TTTTGGCACC AAAATCAACG GGACTTTCCA AAATGTCGTA ATAACCCCGC  
AAAACCGTGG TTTTAGTTGC CCTGAAAGGT TTTACAGCAT TATTGGGGCG

---

951 CCCGTTGACG CAAATGGGCG GTAGGCGTGT ACGGTGGGAG GTCTATATAA  
GGGCAACTGC GTTTACCCGC CATCCGCACA TGCCACCCTC CAGATATATT

---

1001 GCAGAGCTCG TTTAGTGAAC CGTCAGATCG CCTGGAGACG CCATCCACGC  
CGTCTCGAGC AAATCACTTG GCAGTCTAGC GGACCTCTGC GGTAGGTGCG

---

1051 TGTTTTGACC TCCATAGAAG ACACCGGGAC CGATCCAGCC TCCGCGGCCG  
ACAAAACCTGG AGGTATCTTC TGTGGCCCTG GCTAGGTGCG AGGCGCCGGC

---

1101 GGAACGGTGC ATTGGAACGC GGATTCCCCG TGCCAAGAGT GACGTAAGTA  
CCTTGCCACG TAACCTTGCG CTAAGGGGC ACGGTTCTCA CTGCATTCTA

---

1151 CCGCCTATAG ACTCTATAGG CACACCCCTT TGGCTCTTAT GCATGCTATA  
GGCGGATATC TGAGATATCC GTGTGGGGAA ACCGAGAATA CGTACGATAT

---

1201 CTGTTTTTGG CTTGGGGCCT ATACACCCCG GCTCCTTATG CTATAGGTGA  
GACAAAAACC GAACCCCGGA TATGTGGGGG CGAGGAATAC GATATCCACT

---

1251 TGGTATAGCT TAGCCTATAG GTGTGGGTTA TTGACCATTA TTGACCACTC  
ACCATATCGA ATCGGATATC CACACCCAAT AACTGGTAAT AACTGGTGAG

---

1301 CCCTATTGGT GACGATACTT TCCATTACTA ATCCATAACA TGGCTCTTTG  
GGGATAACCA CTGCTATGAA AGGTAATGAT TAGGTATTGT ACCGAGAAAC

---

1351 CCACAACATAT CTCTATTGGC TATATGCCAA TACTCTGTCC TTCAGAGACT  
GGTGTTGATA GAGATAACCG ATATACGGTT ATGAGACAGG AAGTCTCTGA

---

1401 GACACGGACT CTGTATTTTT ACAGGATGGG GTCCATTTAT TATTTACAAA  
CTGTGCCTGA GACATAAAAA TGTCCTACCC CAGGTAAATA ATAAATGTTT

---

1451 TTCACATATA CAACAACGCC GTCCCCCGTG CCCGCAGTTT TTATTAAACA  
AAGTGTATAT GTTGTTCGCG CAGGGGGCAC GGGCGTCAAA AATAATTTGT

---

1501 TAGCGTGGA TCTCCGACAT CTCGGGTACG TGTTCGGAC ATGGGCTCTT  
ATCGCACCTT AGAGGCTGTA GAGCCCATGC ACAAGGCCTG TACCCGAGAA

---

1551 CTCCGGTAGC GGCGGAGCTT CCACATCCGA GCCCTGGTCC CATCCGTCCA  
GAGGCCATCG CCGCCTCGAA GGTGTAGGCT CGGGACCAGG GTAGGCAGGT

---

1601 GCGGCTCATG GTCGCTCGGC AGCTCCTTGC TCCTAACAGT GGAGGCCAGA  
CGCCGAGTAC CAGCGAGCCG TCGAGGAACG AGGATTGTCA CCTCCGGTCT

---

1651 CTTAGGCACA GCACAATGCC CACCACCACC AGTGTGCCGC ACAAGGCCGT  
GAATCCGTGT CGTGTACGG GTGGTGGTGG TCACACGGCG GTTCCGGCA

---

1701 GGCGGTAGGG TATGTGTCTG AAAATGAGCT CGGAGATTGG GCTCGCACCT  
CCGCCATCCC ATACACAGAC TTTTACTCGA GCCTCTAACC CGAGCGTGGA

---

1751 GGACGCAGAT GGAAGACTTA AGGCAGCGGC AGAAGAAGAT GCAGGCAGCT  
CCTGCGTCTA CCTTCTGAAT TCCGTGCGCG TCTTCTCTA CGTCCGTCTA

---

1801 GAGTTGTTGT ATTCTGATAA GAGTCAGAGG TAACTCCCGT TCGGGTGCTG  
CTCAACAACA TAAGACTATT CTCAGTCTCC ATTGAGGGCA ACGCCACGAC

## FIGURE 9 - Page 3

1851 TTAACGGTGG AGGGCAGTGT AGTCTGAGCA GTACTCGTTG CTGCCGCGCG  
AATTGCCACC TCCCGTCACA TCAGACTCGT CATGAGCAAC GACGGCGCGC

---

1901 CGCCACCAGA CATAATAGCT GACAGACTAA CAGACTGTTC CTTTCCATGG  
GCGGTGGTCT GTATTATCGA CTGTCTGATT GTCTGACAAG GAAAGGTACC

---

+2 M A P  
EcoRI  
~~~~~

1951 GTCTTTTCTG CAGTCACCGT CGTCGACCTA AGAATTCACC ATGGCGCCCA
CAGAAAAGAC GTCAGTGGCA GCAGCTGGAT TCTTAAGTGG TACCGCGGGT

+2 I T A Y A Q Q T R G L L G C I I T
2001 TCACGGCGTA CGCCCAGCAG ACAAGGGGCC TCCTAGGGTG CATAATCACC
AGTGCCGCAT GCGGGTCGTC TGTTCCCCGG AGGATCCAC GTATTAGTGG

+2 S L T G R D K N Q V E G E V Q I V
2051 AGCCTAAGT GCGGGGACAA AAACCAAGTG GAGGGTGAGG TCCAGATTGT
TCGGATTGAC GCGCCCTGTT TTTGGTTTAC CTCCCACTCC AGGTCTAACA

+2 S T A A Q T F L A T C I N G V C
2101 GTCAACTGCT GCCCAAACCT TCCTGGCAAC GTGCATCAAT GGGGTGTGCT
CAGTTGACGA CCGGTTTGGA AGGACCGTTG CACGTAGTTA CCCCACACGA

+2 W T V Y H G A G T R T I A S P K G
2151 GGACTGTCTA CCACGGGGCC GGAACGAGGA CCATCGCGTC ACCCAAGGGT
CCTGACAGAT GGTGCCCCGG CTTGCTCTCT GGTAGCGCAG TGGTTTCCCA

+2 P V I Q M Y T N V D Q D L V G W P
2201 CCTGTCTATC AGATGTATAC CAATGTAGAC CAAGACCTTG TGGGCTGGCC
GGACAGTAGG TCTACATATG GTTACATCTG GTTCTGGAAC ACCCGACCGG

+2 A S Q G T R S L T P C T C G S S
2251 CGCTTCGCAA GGTACCCGCT CATTGACACC CTGCACTTGC GGCTCCTCGG
GCGAAGCGTT CCATGGGCGA GTAAGTGTGG GACGTGAACG CCGAGGAGCC

+2 D L Y L V T R H A D V I P V R R R
2301 ACCTTTACCT GGTACAGAGG CACGCCGATG TCATTCCCGT GCGCCGGCGG
TGGAAATGGA CCAGTGCTCC GTGCGGCTAC AGTAAGGGCA CGCGGCCGCC

+2 G D S R G S L L S P R P I S Y L K
2351 GGTGATAGCA GGGGCAGCCT GCTGTGCCCC CGGCCCATTT CCTACTTGAA
CCACTATCGT CCCCCTCGGA CGACAGCGGG GCCGGGTAAA GGATGAACCT

+2 G S S G G P L L C P A G H A V G
2401 AGGCTCCTCG GGGGGTCCGC TGTTGTGCCC CGCGGGGCAC GCCGTGGGCA
TCCGAGGAGC CCCCAGGCG ACAACACGGG GCGCCCCGTG CGGCACCCGT

+2 I F R A A V C T R G V A K A V D F
2451 TATTTAGGGC CGCGGTGTGC ACCCGTGGAG TGGCTAAGGC GGTGGACTTT
ATAAATCCCC GCGCCACACG TGGGCACCTC ACCGATTCCG CCACCTGAAA

+2 I P V E N L E T T M R S P V F T D
2501 ATCCCTGTGG AGAACCTAGA GACAACCATG AGGTCCCCGG TGTTACGGA
TAGGGACACC TCTTGGATCT CTGTTGGTAC TCCAGGGGCC ACAAGTGCCT

PCMV-NS34A
FIGURE 9 - Page 4

+2 N S S P P V V P Q S F Q V A H L
2551 TAACTCCTCT CCACCAGTAG TGCCCCAGAG CTTCCAGGTG GCTCACCTCC
ATTGAGGAGA GGTGGTCATC ACGGGGTCTC GAAGGTCCAC CGAGTGGAGG

+2 H A P T G S G K S T K V P A A Y A
2601 ATGCTCCAC AGGCAGCGGC AAAAGCACCA AGGTCCCGGC TGCATATGCA
TACGAGGGTG TCCGTCGCCG TTTTCGTGGT TCCAGGGCCG ACGTATACGT

+2 A Q G Y K V L V L N P S V A A T L
2651 GCTCAGGGCT ATAAGGTGCT AGTACTCAAC CCCTCTGTTG CTGCAACACT
CGAGTCCCGA TATTCCACGA TCATGAGTTG GGGAGACAAC GACGTTGTGA

+2 G F G A Y M S K A H G I D P N I
2701 GGGCTTTGGT GCTTACATGT CCAAGGCTCA TGGGATCGAT CCTAACATCA
CCCGAAACCA CGAATGTACA GGTTCGAGT ACCCTAGCTA GGATTGTAGT

+2 R T G V R T I T T G S P I T Y S T
2751 GGACCGGGT GAGAACAATT ACCACTGGCA GCCCCATCAC GTACTCCACC
CCTGGCCCCA CTCTTGTTAA TGGTGACCGT CGGGGTAGTG CATGAGGTGG

+2 Y G K F L A D G G C S G G A Y D I
2801 TACGGCAAGT TCCTTGCCGA CGGCGGGTGC TCGGGGGGCG CTTATGACAT
ATGCCGTTCA AGGAACGGCT GCCGCCACG AGCCCCCGC GAATACTGTA

+2 I I C D E C H S T D A T S I L G
2851 AATAATTTGT GACGAGTGCC ACTCCACGGA TGCCACATCC ATCTTGGGCA
TTATTAAACA CTGCTACGG TGAGGTGCCT ACGGTGTAGG TAGAACCCGT

+2 I G T V L D Q A E T A G A R L V V
2901 TTGGCACTGT CCTTGACCAA GCAGAGACTG CGGGGGCGAG ACTGGTTGTG
AACCCTGACA GGAAGTGGT CGTCTCTGAC GCCCCGCTC TGACCAACAC

+2 L A T A T P P G S V T V P H P N I
2951 CTCGCCACCG CCACCCCTCC GGGCTCCGTC ACTGTGCCCC ATCCCAACAT
GAGCGGTGGC GGTGGGAGG CCCGAGGAG TGACACGGGG TAGGGTTGTA

+2 E E V A L S T T G E I P F Y G K
3001 CGAGGAGGTT GCTCTGTCCA CCACCGGAGA GATCCCTTTT TACGGCAAGG
GCTCCTCAA CGAGACAGGT GGTGGCCTCT CTAGGGA AAA ATGCCGTTCC

+2 A I P L E V I K G G R H L I F C H
3051 CTATCCCCT CGAAGTAATC AAGGGGGGA GACATCTCAT CTTCTGTCAT
GATAGGGGA GCTTCATTAG TTCCCCCT CTGTAGAGTA GAAGACAGTA

+2 S K K K C D E L A A K L V A L G I
3101 TCAAAGAAGA AGTGCGACGA ACTCGCCGCA AAGCTGGTCG CATTGGGCAT
AGTTTCTTCT TCACGCTGCT TGAGCGGCGT TTCGACCAGC GTAACCCGTA

+2 N A V A Y Y R G L D V S V I P T
3151 CAATGCCGTG GCCTACTACC GCGGTCTTGA CGTGTCCGTC ATCCCGACCA
GTTACGGCAC CGGATGATGG CGCCAGAACT GCACAGGCAG TAGGGCTGGT

+2 S G D V V V V A T D A L M T G Y T
3201 GCGGCGATGT TGTCGTCGTG GCAACCGATG CCCTCATGAC CGGCTATACC
CGCCGCTACA ACAGCAGCAC CGTTGGCTAC GGGAGTACTG GCCGATATGG

002277-647450

pCMV-NS34A

FIGURE 9 - Page 5

+2 G D F D S V I D C N T C V T Q T V
3251 GCGACTTCG ACTCGGTGAT AGACTGCAAT ACGTGTGTCA CCCAGACAGT
CCGCTGAAGC TGAGCCACTA TCTGACGTTA TGCACACAGT GGGTCTGTCA

+2 D F S L D P T F T I E T I T L P
3301 CGATTTTCAGC CTTGACCCTA CTTTACCAT TGAGACAATC ACGTCCCCC
GCTAAAGTCG GAACTGGGAT GGAAGTGGTA ACTCTGTTAG TGCGAGGGGG

+2 Q D A V S R T Q R R G R T G R G K
3351 AAGATGCTGT CTCCCGCACT CAACGTCGGG GCAGGACTGG CAGGGGGAAG
TTCTACGACA GAGGGCGTGA GTTGACGCCC CGTCCTGACC GTCCCCCTC

+2 P G I Y R F V A P G E R P S G M F
3401 CCAGGCATCT ACAGATTTGT GGCACCGGGG GAGCGCCCCT CCGGCATGTT
GTCCCGTAGA TGTCTAAACA CCGTGGCCCC CTCGCGGGGA GGCCGTACAA

+2 D S S V L C E C Y D A G C A W Y
3451 CGACTCGTCC GTCCTCTGTG AGTGCTATGA CGCAGGCTGT GCTTGGTATG
GCTGAGCAGG CAGGAGACAC TCACGATACT GCGTCCGACA CGAACCATAC

+2 E L T P A E T T V R L R A Y M N T
3501 AGCTCACGCC CGCCGAGACT ACAGTTAGGC TACGAGCGTA CATGAACACC
TCGAGTGCGG GCGGCTCTGA TGTCAATCCG ATGCTCGCAT GTACTTGTGG

+2 P G L P V C Q D H L E F W E G V F
3551 CCGGGGCTTC CCGTGTGCCA GGACCATCTT GAATTTTGGG AGGGCGTCTT
GGCCCCGAAG GGCACACGGT CCTGGTAGAA CTAAAAACCC TCCCGCAGAA

+2 T G L T H I D A H F L S Q T K Q
StuI
~~~~~  
3601 TACAGGCCTC ACTCATATAG ATGCCCACTT TCTATCCCAG ACAAAGCAGA  
ATGTCCGGAG TGAGTATATC TACGGGTGAA AGATAGGGTC TGTTTCGTCT

+2 S G E N L P Y L V A Y Q A T V C A  
3651 GTGGGGAGAA CCTTCCTTAC CTGGTAGCGT ACCAAGCCAC CGTGTGCGCT  
CACCCTCTT GGAAGGAATG GACCATCGCA TGGTTCGGTG GCACACGCGA

+2 R A Q A P P P S W D Q M W K C L I  
3701 AGGGCTCAAG CCCCTCCCCC ATCGTGGGAC CAGATGTGGA AGTGTGTTGAT  
TCCCGAGTTC GGGGAGGGGG TAGCACCTTG GTCTACACCT TCACAACTA

+2 R L K P T L H G P T P L L Y R L  
3751 TCGCCTCAAG CCCACCCTCC ATGGGCCAAC ACCCCTGCTA TACAGACTGG  
AGCGGAGTTC GGGTGGGAGG TACCCGGTTG TGGGGACGAT ATGTCTGACC

+2 G A V Q N E I T L T H P V T K Y I  
3801 GCGCTGTTCA GAATGAAATC ACCCTGACGC ACCCAGTCAC CAAATACATC  
CGCGACAAGT CTTACTTTAG TGGGACTGCG TGGGTCACTG GTTTATGTAG

+2 M T C M S A D L E V V T S T W V L  
3851 ATGACATGCA TGTCGGCCGA CCTGGAGGTC GTCACGAGCA CCTGGGTGCT  
TACTGTACGT ACAGCCGGCT GGACCTCCAG CAGTGCTCGT GGACCCACGA

+2 V G G V L A A L A A Y C L S T G  
3901 CGTTGGCGGC GTCCTGGCTG CTTTGGCCGC GTATTGCCTG TCAACAGGCT  
GCAACCGCCG CAGGACCGAC GAAACCGGCG CATAACGGAC AGTTGTCCGA

pCMV-NS34A  
**FIGURE 9 - Page 6**

```

+2 C V V I V G R V V L S G K P A I I
3951 GCGTGGTCAT AGTGGGCAGG GTCGTCTTGT CCGGGAAGCC GGCAATCATA
      CGCACCAGTA TCACCCGTCC CAGCAGAACA GGCCCTTCGG CCGTTAGTAT

+2 P D R E V L Y R E F D E M E E C
4001 CCTGACAGGG AAGTCCTCTA CCGAGAGTTC GATGAGATGG AAGAGTGCTA
      GGACTGTCCC TTCAGGAGAT GGCTCTCAAG CTA CTCTCTACC TTCTCACGAT

      BamHI      MluI
      ~~~~~
4051 GGATCCACTA CGCGTTAGAG CTCGCTGATC AGCCTCGACT GTGCCTTCTA
 CCTAGGTGAT GCGCAATCTC GAGCGACTAG TCGGAGCTGA CACGGAAGAT

4101 GTTGCCAGCC ATCTGTTGTT TGCCCCCTCCC CCGTGCCTTC CTTGACCCTG
 CAACGGTCGG TAGACAACAA ACGGGGAGGG GGCACGGAAG GAACTGGGAC

4151 GAAGGTGCCA CTCCCCTGT CTTTCTCTAA TAAATGAGG AAATTGCATC
 CTTCCACGGT GAGGGTGACA GGAAAGGATT ATTTTACTCC TTTAACGTAG

4201 GCATTGTCTG AGTAGGTGTC ATTCTATTCT GGGGGGTGGG GTGGGGCAGG
 CGTAACAGAC TCATCCACAG TAAGATAAGA CCCCCACCC CACCCCGTCC

4251 ACAGCAAGGG GGAGGATTGG GAAGACAATA GCAGGCATGC TGGGGAGCTC
 TGTCGTTCCC CCTCCTAACC CTTCTGTTAT CGTCCGTACG ACCCCTCGAG

4301 TTCCGCTTCC TCGCTCACTG ACTCGCTGCG CTCGGTCGTT CGGCTGCGGC
 AAGGCGAAGG AGCGAGTGAC TGAGCGACGC GAGCCAGCAA GCCGACGCCG

4351 GAGCGGTATC AGCTCACTCA AAGGCGGTAA TACGGTTATC CACAGAATCA
 CTCGCCATAG TCGAGTGAGT TTCCGCCATT ATGCCAATAG GTGTCTTAGT

4401 GGGGATAACG CAGGAAAGAA CATGTGAGCA AAAGGCCAGC AAAAGGCCAG
 CCCCTATTGC GTCCTTTCTT GTACACTCGT TTTCCGGTCG TTTTCCGGTC

4451 GAACCGTAAA AAGGCCGCGT TGCTGGCGTT TTTCCATAGG CTCCGCCCCC
 CTTGGCATT TTCCGGCGCA ACGACCGCAA AAAGGTATCC GAGGCGGGGG

4501 CTGACGAGCA TCACAAAAAT CGACGCTCAA GTCAGAGGTG GCGAAACCCG
 GACTGCTCGT AGTGTTTTTA GCTGCGAGTT CAGTCTCCAC CGCTTTGGGC

4551 ACAGGACTAT AAAGATACCA GGCCTTTCCC CCTGGAAGCT CCCTCGTGCG
 TGTCCTGATA TTTCTATGGT CCGCAAAGGG GGACCTTCGA GGGAGCACGC

4601 CTCTCCTGTT CCGACCCTGC CGCTTACCGG ATACCTGTCC GCCTTTCTCC
 GAGAGGACAA GGCTGGGACG GCGAATGGCC TATGGACAGG CGGAAAGAGG

4651 CTTCCGGGAAG CGTGGCGCTT TCTCAATGCT CACGCTGTAG GTATCTCAGT
 GAAGCCCTTC GCACCGCGAA AGAGTTACGA GTGCGACATC CATAGAGTCA

4701 TCGGTGTAGG TCGTTCGCTC CAAGCTGGGC TGTGTGCACG AACCCCCCGT
 AGCCACATCC AGCAAGCGAG GTTCGACCCG ACACACGTGC TTGGGGGGCA

4751 TCAGCCCGAC CGCTGCGCCT TATCCGGTAA CTATCGTCTT GAGTCCAACC
 AGTCGGGCTG GCGACGCGGA ATAGGCCATT GATAGCAGAA CTCAGGTTGG

4801 CGGTAAGACA CGACTTATCG CCACTGGCAG CAGCCACTGG TAACAGGATT
 GCCATTCTGT GCTGAATAGC GGTGACCGTC GTCGGTGACC ATTGTCCTAA

```

Downloaded from www.sciencedirect.com

## FIGURE 9 - Page 7

4851 AGCAGAGCGA GGTATGTAGG CGGTGCTACA GAGTTCTTGA AGTGGTGGCC  
TCGTCTCGCT CCATACATCC GCCACGATGT CTCAAGAACT TCACCACCGG

---

4901 TAACTACGGC TACACTAGAA GGACAGTATT TGGTATCTGC GCTCTGCTGA  
ATTGATGCCG ATGTGATCTT CCTGTCATAA ACCATAGACG CGAGACGACT

---

4951 AGCCAGTTAC CTTCGGAAAA AGAGTTGGTA GCTCTTGATC CGGCAAACAA  
TCGGTCAATG GAAGCCTTTT TCTCAACCAT CGAGAACTAG GCCGTTTGTT

---

5001 ACCACCGCTG GTAGCGGTGG TTTTTTTGTT TGCAAGCAGC AGATTACGCG  
TGGTGGCGAC CATCGCCACC AAAAAAACA ACGTTCGTCG TCTAATGCGC

---

5051 CAGAAAAAAA GGATCTCAAG AAGATCCTTT GATCTTTTCT ACGGGGTCTG  
GTCTTTTTTT CCTAGAGTTC TTCTAGGAAA CTAGAAAAGA TGCCCCAGAC

---

5101 ACGCTCAGTG GAACGAAAAC TCACGTTAAG GGATTTTGGT CATGAGATTA  
TGCGAGTCAC CTTGCTTTTG AGTGCAATTC CCTAAAACCA GTACTCTAAT

---

5151 TCAAAAAGGA TCTTCACCTA GATCCTTTTA AATTAAAAAT GAAGTTTTAA  
AGTTTTTCCT AGAAGTGGAT CTAGGAAAAT TTAATTTTAA CTTCAAAATT

---

5201 ATCAATCTAA AGTATATATG AGTAAACTTG GTCTGACAGT TACCAATGCT  
TAGTTAGATT TCATATATAC TCATTTGAAC CAGACTGTCA ATGGTTACGA

---

5251 TAATCAGTGA GGCACCTATC TCAGCGATCT GTCTATTTTG TTCATCCATA  
ATTAGTCACT CCGTGGATAG AGTCGCTAGA CAGATAAAGC AAGTAGGTAT

---

5301 GTTGCCTGAC TCCCCGTCGT GTAGATAACT ACGATACGGG AGGGCTTACC  
CAACGGACTG AGGGGCAGCA CATCTATTGA TGCTATGCCC TCCCGAATGG

---

5351 ATCTGGCCCC AGTGCTGCAA TGATACCGCG AGACCCACGC TCACCGGCTC  
TAGACCGGGG TCACGACGTT ACTATGGCGC TCTGGGTGCG AGTGGCCGAG

---

5401 CAGATTTATC AGCAATAAAC CAGCCAGCCG GAAGGGCCGA GCGCAGAAGT  
GTCTAAATAG TCGTTATTTG GTCGGTCGGC CTTCCCGGCT CGCGTCTTCA

---

5451 GGTCTGCAA CTTTATCCGC CTCCATCCAG TCTATTAATT GTTGCCGGGA  
CCAGGACGTT GAAATAGGCG GAGGTAGGTC AGATAATTAA CAACGGCCCT

---

5501 AGCTAGAGTA AGTAGTTCGC CAGTTAATAG TTTGCGCAAC GTTGTTGCCA  
TCGATCTCAT TCATCAAGCG GTCAATTATC AAACGCGTTG CAACAACGGT

---

5551 TTGCTACAGG CATCGTGGTG TCACGCTCGT CGTTTGGTAT GGCTTCATT  
AACGATGTCC GTAGCACCAC AGTGCGAGCA GCAAACCATA CCGAAGTAAG

---

5601 AGCTCCGGTT CCCAACGATC AAGGCGAGTT ACATGATCCC CCATGTTGTG  
TCGAGGCCAA GGGTTGCTAG TTCCGCTCAA TGTACTAGG GGTACAACAC

---

5651 CAAAAAAGCG GTTAGCTCCT TCGGTCCTCC GATCGTTGTC AGAAGTAAGT  
GTTTTTTCGC CAATCGAGGA AGCCAGGAGG CTAGCAACAG TCTTCATTCA

---

5701 TGGCCGCACT GTTATCACTC ATGGTTATGG CAGCACTGCA TAATTCTCTT  
ACCGGCGTCA CAATAGTGAG TACCAATACC GTCGTGACGT ATTAAGAGAA

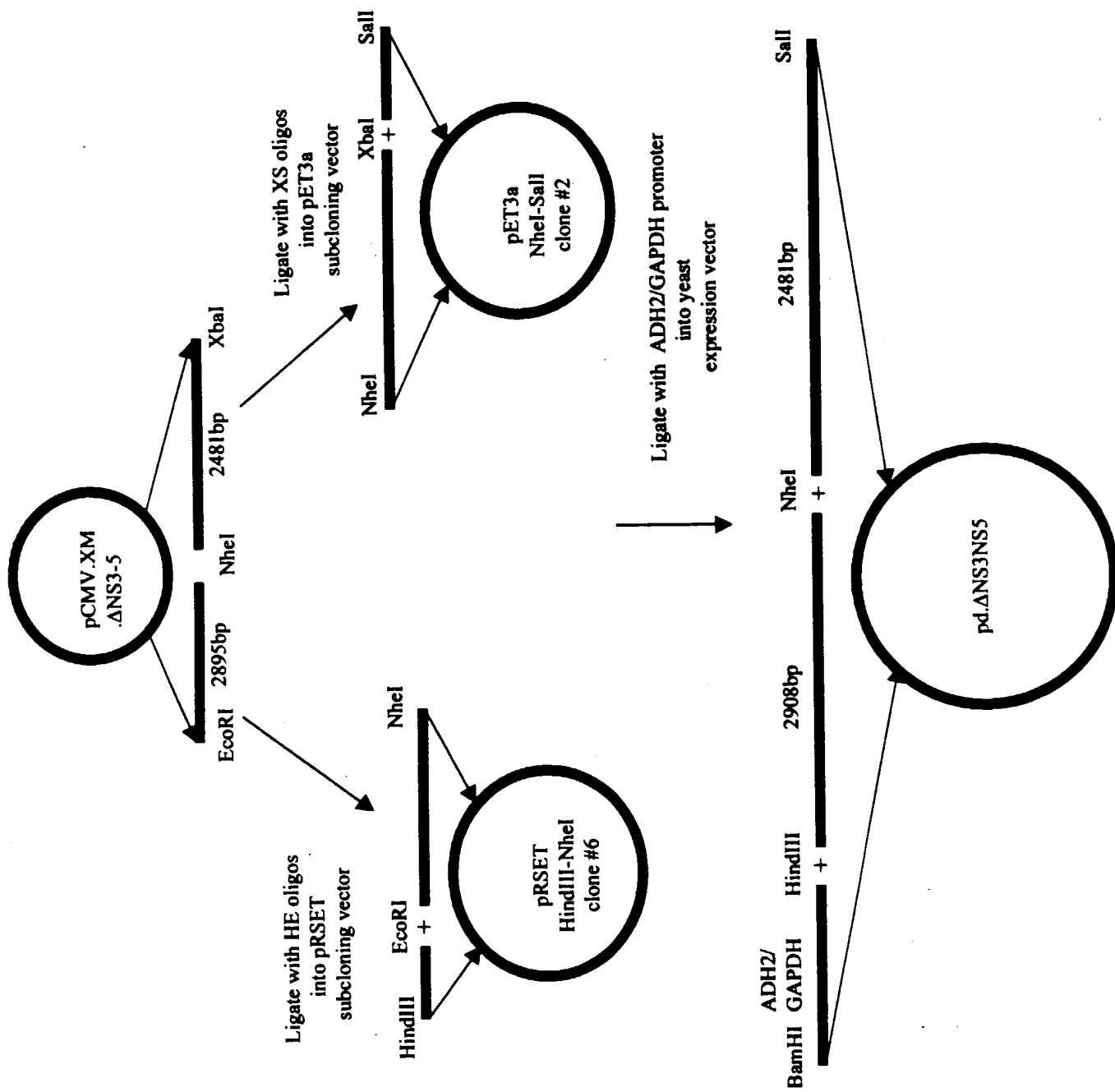
---

5751 ACTGTCATGC CATCCGTAAG ATGCTTTTCT GTGACTGGTG AGTACTCAAC  
TGACAGTACG GTAGGCATT CACTGACCAC TCATGAGTTG

|      |                           |                          |                           |                          |                          |
|------|---------------------------|--------------------------|---------------------------|--------------------------|--------------------------|
| 5801 | CAAGTCATTC<br>GTTTCAGTAAG | TGAGAATAGT<br>ACTCTTATCA | GTTATGCGGCG<br>CATACGCCGC | ACCGAGTTGC<br>TGGCTCAACG | TCTTGCCCCG<br>AGAACGGGCC |
| 5851 | CGTCAATACG<br>GCAGTTATGC  | GGATAATACC<br>CCTATTATGG | GCGCCACATA<br>CGCGGTGTAT  | GCAGAACTTT<br>CGTCTTGAAA | AAAAGTGCTC<br>TTTTCACGAG |
| 5901 | ATCATTGGAA<br>TAGTAACCTT  | AACGTTCTTC<br>TTGCAAGAAG | GGGGCGAAAA<br>CCCCGCTTTT  | CTCTCAAGGA<br>GAGAGTTCCT | TCTTACCGCT<br>AGAATGGCGA |
| 5951 | GTTGAGATCC<br>CAACTCTAGG  | AGTTCGATGT<br>TCAAGCTACA | AACCCACTCG<br>TTGGGTGAGC  | TGCACCCAAC<br>ACGTGGGTG  | TGATCTTCAG<br>ACTAGAAGTC |
| 6001 | CATCTTTTAC<br>GTAGAAAATG  | TTTCACCAGC<br>AAAGTGGTGC | GTTTCTGGGT<br>CAAAGACCCA  | GAGCAAAAAC<br>CTCGTTTTTG | AGGAAGGCAA<br>TCCTTCCGTT |
| 6051 | AATGCCGCAA<br>TTACGGCGTT  | AAAAGGGAAT<br>TTTTCCCTTA | AAGGGCGACA<br>TTCCCGCTGT  | CGGAAATGTT<br>GCCTTTACAA | GAATACTCAT<br>CTTATGAGTA |
| 6101 | ACTCTTCCTT<br>TGAGAAGGAA  | TTTCAATATT<br>AAAGTTATAA | ATTGAAGCAT<br>TAACTTCGTA  | TTATCAGGGT<br>AATAGTCCCA | TATTGTCTCA<br>ATAACAGAGT |
| 6151 | TGAGCGGATA<br>ACTCGCCTAT  | CATATTTGAA<br>GTATAAACTT | TGTATTTAGA<br>ACATAAATCT  | AAAATAAACA<br>TTTTATTGT  | AATAGGGGTT<br>TTATCCCCAA |
| 6201 | CCGCGCACAT<br>GGCGCGTGTA  | TTCCCCGAAA<br>AAGGGGCTTT | AGTGCCACCT<br>TCACGGTGGA  | GACGTCTAAG<br>CTGCAGATTC | AAACCATTAT<br>TTTGGAATA  |
| 6251 | TATCATGACA<br>ATAGTACTGT  | TTAACTATA<br>AATTGGATAT  | AAAAATAGGCG<br>TTTTATCCGC | TATCACGAGG<br>ATAGTGCTCC | CCCTTTCGTC<br>GGGAAAGCAG |



Diagram 1



# FIGURE 11 - Page 1

MetAlaAlaTyrAlaAlaGlnGlyTyrLysValLeuVal  
 2 AGCTTACAAAACAAATTCACCATGGCTGCATATGCAGCTCAGGGCTATAAGGTGCTAGTA  
 TCGAATGTTTTGTTTAAGTGGTACCGACGTATACGTCGAGTCCCGATATTCCACGATCAT  
 ^ ^ ^ ^  
 1 HIND3, 21 NCOI, 30 NDEI, 58 SCAI,

LeuAsnProSerValAlaAlaThrLeuGlyPheGlyAlaTyrMetSerLysAlaHisGly  
 62 CTCAACCCCTCTGTTGCTGCAACACTGGGCTTTGGTGCTTACATGTCCAAGGCTCATGGG  
 GAGTTGGGGAGACAACGACGTTGTGACCCGAAACCACGAATGTACAGGTTCCGAGTACCC

IleAspProAsnIleArgThrGlyValArgThrIleThrThrGlySerProIleThrTyr  
 122 ATCGATCCTAACATCAGGACCGGGGTGAGAACAATTACCACTGGCAGCCCCATCACGTAC  
 TAGCTAGGATTGTAGTCCTGGCCCCACTCTTGTTAATGGTGACCGTCGGGGTAGTGCATG  
 ^  
 122 CLAI,

SerThrTyrGlyLysPheLeuAlaAspGlyGlyCysSerGlyGlyAlaTyrAspIleIle  
 182 TCCACCTACGGCAAGTTCCTTGCCGACGGCGGGTGCTCGGGGGGCGCTTATGACATAATA  
 AGGTGGATGCCGTTCAAGGAACGGCTGCCGCCACGAGCCCCCGGAATACTGTATTAT

IleCysAspGluCysHisSerThrAspAlaThrSerIleLeuGlyIleGlyThrValLeu  
 242 ATTTGTGACGAGTGCCACTCCACGGATGCCACATCCATCTTGGGCATTGGCACTGTCCTT  
 TAAACACTGCTCACGGTGAGGTGCCTACGGTGTAGGTAGAACCCGTAACCGTGACAGGAA

AspGlnAlaGluThrAlaGlyAlaArgLeuValValLeuAlaThrAlaThrProProGly  
 302 GACCAAGCAGAGACTGCGGGGCGAGACTGGTTGTGCTCGCCACCGCCACCCCTCCGGGC  
 CTGGTTCGTCTCTGACGCCCCCGCTCTGACCAACACGAGCGGTGGCGGTGGGGAGGCCCC  
 ^  
 309 ALWN1,

SerValThrValProHisProAsnIleGluGluValAlaLeuSerThrThrGlyGluIle  
 362 TCCGTCACTGTGCCCCATCCCAACATCGAGGAGGTGCTCTGTCCACCACCGGAGAGATC  
 AGGCAGTGACACGGGGTAGGGTTGTAGCTCCTCCAACGAGACAGGTGGTGGCCTCTCTAG

ProPheTyrGlyLysAlaIleProLeuGluValIleLysGlyGlyArgHisLeuIlePhe  
 422 CCTTTTACGGCAAGGCTATCCCCCTCGAAGTAATCAAGGGGGGAGACATCTCATCTTC  
 GGAAAAATGCCGTTCCGATAGGGGGAGCTTCATTAGTTCCCCCCTCTGTAGAGTAGAAG

CysHisSerLysLysLysCysAspGluLeuAlaAlaLysLeuValAlaLeuGlyIleAsn  
 482 TGTCATTCAAAGAAGAAGTGCGACGAACTCGCCGCAAAGCTGGTCGCATTGGGCATCAAT  
 ACAGTAAGTTTCTTCTTCACGCTGCTTGAGCGGCGTTTCGACCAGCGTAACCCGTAGTTA

AlaValAlaTyrTyrArgGlyLeuAspValSerValIleProThrSerGlyAspValVal  
 542 GCCGTGGCCTACTACCGCGGTCTTGACGTGTCCGTCATCCCGACCGGCGATGTTGTC  
 CGGCACCGGATGATGGCGCCAGAACTGCACAGGCAGTAGGGCTGGTCGCCGCTACAACAG  
 ^ ^  
 556 SAC2, 566 DRD1,

ValValAlaThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAsp  
 602 GTCGTGGCAACCGATGCCCTCATGACGGGCTATACCGGCGACTTCGACTCGGTGATAGAC  
 CAGCACCGTTGGCTACGGGAGTACTGGCCGATATGGCCGCTGAAGCTGAGCCACTATCTG  
 ^  
 621 BSPH1,

CysAsnThrCysValThrGlnThrValAspPheSerLeuAspProThrPheThrIleGlu

562 TGCAATACGTGTGTACCCAGACAGTCGATTTTCAGCCTTGACCCTACCTTCACCATTGAG  
ACGTTATGCACACAGTGGGTCTGTCTCAGCTAAAGTCGGAAGTGGGATGGAAGTGGTAACTC

ThrIleThrLeuProGlnAspAlaValSerArgThrGlnArgArgGlyArgThrGlyArg  
722 ACAATCACGCTCCCCAAGATGCTGTCTCCCGCACTCAACGTCGGGGCAGGACTGGCAGG  
TGTTAGTGCGAGGGGGTTCTACGACAGAGGGCGTGAGTTGCAGCCCCGTCTGACCGTCC

GlyLysProGlyIleTyrArgPheValAlaProGlyGluArgProSerGlyMetPheAsp  
782 GGGAAAGCCAGGCATCTACAGATTTGTGGCACCGGGGAGCGCCCCCTCCGGCATGTTTCGAC  
CCCTTCGGTCCGTAGATGTCTAAACACCGTGGCCCCCTCGCGGGGAGGCCGTACAAGCTG

822 BGLI, 839 DRD1,

SerSerValLeuCysGluCysTyrAspAlaGlyCysAlaTrpTyrGluLeuThrProAla  
842 TCGTCCGTCTCTGTGAGTGCTATGACGCAGGCTGTGCTTGGTATGAGCTCACGCCCGCC  
AGCAGGCAGGAGACACTCACGATACTGCGTCCGACACGAACCATACTCGAGTGGGGGCGG

887 SACI,

GluThrThrValArgLeuArgAlaTyrMetAsnThrProGlyLeuProValCysGlnAsp  
902 GAGACTACAGTTAGGCTACGAGCGTACATGAACACCCCGGGGCTTCCCGTGTGCCAGGAC  
CTCTGATGTCAATCCGATGCTCGCATGTACTTGTGGGGCCCCGAAGGGCACACGGTCTCTG

937 SMAI XMAI,

HisLeuGluPheTrpGluGlyValPheThrGlyLeuThrHisIleAspAlaHisPheLeu  
962 CATCTTGAATTTTGGGAGGGCGTCTTTACAGGCCTCACTCATATAGATGCCCACTTTCTA  
GTAGAACTTAAAACCCTCCCGCAGAAATGTCCGGAGTGAGTATATCTACGGGTGAAAGAT

991 STUI,

SerGlnThrLysGlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGlnAlaThrVal  
1022 TCCCAGACAAAGCAGAGTGGGGAGAACCTTCCCTACCTGGTAGCGTACCAAGCCACCGTG  
AGGGTCTGTTTTCGTCTCACCCCTCTTGAAGGAATGGACCATCGCATGGTTCGGTGGCAC

1075 DRA3,

CysAlaArgAlaGlnAlaProProProSerTrpAspGlnMetTrpLysCysLeuIleArg  
1082 TGCCTAGGGCTCAAGCCCCCTCCCCATCGTGGGACCAGATGTGGAAGTGTGTGATTGCG  
ACCGCATCCCGAGTTCCGGGGAGGGGGTAGCACCCCTGGTCTACACCTTCACAACTAAGCG

LeuLysProThrLeuHisGlyProThrProLeuLeuTyrArgLeuGlyAlaValGlnAsn  
1142 CTCAAGCCCACCCTCCATGGGCCAACACCCCTGCTATACAGACTGGGCGCTGTTTCAAGAT  
GAGTTCCGGTGGGAGGTACCCGGTTGTGGGGACGATATGTCTGACCCGCGACAAGTCTTA

1156 NCOI,

GluIleThrLeuThrHisProValThrLysTyrIleMetThrCysMetSerAlaAspLeu  
1202 GAAATCACCTGACGCACCCAGTCACCAAATACATCATGACATGCATGTGGCCGACCTG  
CTTTAGTGGGACTGCGTGGGTGCTGAGTGTGAGTGTACGTACAGCCGGCTGGAC

1236 BSPH1, 1240 DRD1, 1243 AVA3, 1251 EAG1 XMA3, 1256 DRD1,

GluValValThrSerThrTrpValLeuValGlyGlyValLeuAlaAlaLeuAlaAlaTyr  
1262 GAGGTCGTACAGACACCTGGGTGCTCGTTGGCGGCGTCTGGCTGCTTTGGCCGCGTAT  
CTCCAGCAGTGCTCGTGGACCCACGAGCAACCGCCGAGGACCGACGAACCGGCGCATA

# FIGURE 11 - Page 3

1322 CysLeuSerThrGlyCysValValIleValGlyArgValValLeuSerGlyLysProAla  
 TGCCTGTCAACAGGCTGCGTGGTTCATAGTGGGCAGGGTCGTCTTGTCCGGGAAGCCGGCA  
 ACGGACAGTTGTCCGACGCACCAGTATCACCCGTCCCAGCAGAACAGGCCCTTCGGCCGT  
 1375 NAEI,  
 1382 IleIleProAspArgGluValLeuTyrArgGluPheAspGluMetGluGluCysSerGln  
 ATCATACCTGACAGGGAAGTCCTCTACCGAGAGTTCGATGAGATGGAAGAGTGCTCTCAG  
 TAGTATGGACTGTCCCTTCAGGAGATGGCTCTCAAGCTACTCTACCTTCTCACGAGAGTC  
 1391 DRD1,  
 1442 HisLeuProTyrIleGluGlnGlyMetMetLeuAlaGluGlnPheLysGlnLysAlaLeu  
 CACTTACCGTACATCGAGCAAGGGATGATGCTCGCCGAGCAGTTCAAGCAGAAGGCCCTC  
 GTGAATGGCATGTAGCTCGTTCCTACTACGAGCGGCTCGTCAAGTTCGTCTTCCGGGAG  
 1502 GlyLeuLeuGlnThrAlaSerArgGlnAlaGluValIleAlaProAlaValGlnThrAsn  
 GGCCTCCTGCAGACCGCGTCCCGTCAGGCAGAGGTTATCGCCCCTGTGTCCAGACCAAC  
 CCGGAGGACGTCTGGCGCAGGGCAGTCCGTCTCCAATAGCGGGGACGACAGGTCTGGTTG  
 1508 PSTI, 1513 TTH3I,  
 1562 TrpGlnLysLeuGluThrPheTrpAlaLysHisMetTrpAsnPheIleSerGlyIleGln  
 TGGCAAAAACCTCGAGACCTTCTGGGCGAAGCATATGTGGAAGTTCATCAGTGGGATACAA  
 ACCGTTTTTGAGCTCTGGAAGACCCGCTTCGTATACACCTTGAAGTAGTCACCCTATGTT  
 1571 XHOI, 1592 NDEI,  
 1622 TyrLeuAlaGlyLeuSerThrLeuProGlyAsnProAlaIleAlaSerLeuMetAlaPhe  
 TACTTGGCGGGCTTGTC AACGCTGCCTGGTAACCCCGCCATTGCTTCATTGATGGCTTTT  
 ATGAACCGCCCCAACAGTTGCGACGGACATTGGGGCGGTAACGAAGTAACCTACCGAAAA  
 1649 BSTE2,  
 1682 ThrAlaAlaValThrSerProLeuThrThrSerGlnThrLeuLeuPheAsnIleLeuGly  
 ACAGCTGCTGTCAACGCTGCTGCTGGTAACCCCGCCATTGCTTCATTGATGGCTTTT  
 TGTGACGACAGTGGTTCGGGTGATTGGTGATCGGTTTGGGAGGAGAAGTTGTATAACCCC  
 1683 ALWN1 PVU2,  
 1742 GlyTrpValAlaAlaGlnLeuAlaAlaProGlyAlaAlaThrAlaPheValGlyAlaGly  
 GGGTGGGTGGCTGCCCAGCTCGCCGCCCCCGGTGCCGCTACTGCCTTTGTGGGCGCTGGC  
 CCCACCCACCGACGGGTCGAGCGGGCGGGGCCACGGCGATGACGGAAACACCCGCGACCG  
 1800 ESP1,  
 1802 LeuAlaGlyAlaAlaIleGlySerValGlyLeuGlyLysValLeuIleAspIleLeuAla  
 TTAGCTGGCGCCGCGCATCGGCAGTGTGGACTGGGGAAGGTCCTCATAGACATCCTTGCA  
 AATCGACCGCGCGGTAGCCGTACAACCTGACCCCTTCCAGGAGTATCTGTAGGAACGT  
 1808 KAS1 NARI,  
 1862 GlyTyrGlyAlaGlyValAlaGlyAlaLeuValAlaPheLysIleMetSerGlyGluVal  
 GGGTATGGCGCGGGCGTGGCGGGAGCTCTTGTGGCATTCAAGATCATGACGGGTGAGGTC  
 CCCATACCGCGCCCGCACCGCCCTCGAGAACACCGTAAGTTCTAGTACTCGCCACTCCAG

00327-64460

# FIGURE 11 - Page 4

1884 SACI, 1905 BSPH1,

1922 ProSerThrGluAspLeuValAsnLeuLeuProAlaIleLeuSerProGlyAlaLeuVal  
 CCCTCCACGGAGGACCTGGTCAATCTACTGCCCCCATCTCTCGCCCGGAGCCCTCGTA  
 GGGAGGTGCCTCCTGGACCAGTTAGATGACGGGCGGTAGGAGAGCGGGCCTCGGGAGCAT  
 ^

1934 TTH3I,

1982 ValGlyValValCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaVal  
 GTCGGCGTGGTCTGTGCAGCAATACTGCGCCGGCACGTTGGCCCGGGCGAGGGGCGAGTG  
 CAGCCGCACACAGACAGTCGTTATGACGCGGGCGGTGCAACCGGGCCGCTCCCCCGTCAC  
 ^

2010 NAEI, 2023 SMAI XMAI,

2042 GlnTrpMetAsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValSerProThrHis  
 CAGTGGATGAACCGGCTGATAGCCTTCGCCTCCCGGGGAACCATGTTTCCCCACGCAC  
 GTCACCTACTTGGCCGACTATCGGAAGCGGAGGGCCCCCTTGGTACAAAGGGGGTGCCTG  
 ^

2073 SMAI XMAI, 2099 DRA3,

2102 TyrValProGluSerAspAlaAlaAlaArgValThrAlaIleLeuSerSerLeuThrVal  
 TACGTGCCGGAGAGCGATGCAGCTGCCCCGCTCACTGCCATACTCAGCAGCCTCACTGTA  
 ATGCACGGCCTCTCGCTACGTGACGGGCGCAGTGACGGTATGAGTCGTCGGAGTGACAT  
 ^

2121 PVU2,

2162 ThrGlnLeuLeuArgArgLeuHisGlnTrpIleSerSerGluCysThrThrProCysSer  
 ACCCAGCTCCTGAGGCGACTGCACCACTGGGATAAGCTCGGAGTGTAACCACTCCATGCTCC  
 TGGGTGCGAGGACTCCGCTGACGTGGTCACCTATTCGAGCCTCACATGGTGAGGTACGAGG  
 ^

2165 ALWN1, 2170 MST2,

2222 GlySerTrpLeuArgAspIleTrpAspTrpIleCysGluValLeuSerAspPheLysThr  
 GGTTCCTGGCTAAGGGACATCTGGGACTGGATATGCGAGGTGTGAGCGACTTTAAGACC  
 CCAAGGACCGATTCCCTGTAGACCCTGACCTATACGCTCCCAACTCGCTGAAATTCTGG  
 ^

2226 ECON1,

2282 TrpLeuLysAlaLysLeuMetProGlnLeuProGlyIleProPheValSerCysGlnArg  
 TGGCTAAAAGCTAAGCTCATGCCACAGCTGCCTGGGATCCCCTTTGTGTCCTGCCAGCGC  
 ACCGATTTTTCGATTGAGTACGGTGTGACGGACCCTAGGGGAAACACAGGACGGTCGCG  
 ^

2291 ESP1, 2306 PVU2, 2316 BAMHI,

2342 GlyTyrLysGlyValTrpArgGlyAspGlyIleMetHisThrArgCysHisCysGlyAla  
 GGGTATAAGGGGGTCTGGCGAGGGGACGGCATCATGCACACTCGCTGCCACTGTGGAGCT  
 CCCATATTCCCCAGACCGCTCCCCTGCCGTAGTACGTGTGAGCGACGGTGACACCTCGA

2402 GluIleThrGlyHisValLysAsnGlyThrMetArgIleValGlyProArgThrCysArg  
 GAGATCACTGGACATGTCAAAAACGGGACGATGAGGATCGTCGGTCCTAGGACCTGCAGG  
 CTCTAGTGACCTGTACAGTTTTTGCCCTGCTACTCCTAGCAGCCAGGATCCTGGACGTCC  
 ^

2431 BSAB1, 2447 AVR2, 2454 SSE83871, 2455 PSTI,

2462 AsnMetTrpSerGlyThrPheProIleAsnAlaTyrThrThrGlyProCysThrProLeu  
 AACATGTGGAGTGGGACCTTCCCCATTAATGCCTACACCACGGGGCCCTGTACCCCCCTT  
 TTGTACACCTCACCTGGAAGGGGTAATTACGGATGTGGTGCCCCGGGGACATGGGGGGAA  
 ^

2486 ASE1, 2503 APAI,

2522 ProAlaProAsnTyrThrPheAlaLeuTrpArgValSerAlaGluGluTyrValGluIle  
CCTGCGCCGAACCTACACGTTTCGCGCTATGGAGGGTGTCTGCAGAGGAATACGTGGAGATA  
GGACGCGGCTTGATGTGCAAGCGCGATACCTCCACAGACGTCTCCTTATGCACCTCTAT

2559 PSTI,

2582 ArgGlnValGlyAspPheHisTyrValThrGlyMetThrThrAspAsnLeuLysCysPro  
AGGCAGGTGGGGGACTTCCACTACGTGACGGGTATGACTACTGACAATCTTAAATGCCCCG  
TCCGTCCACCCCCTGAAGGTGATGCACTGCCCATACTGATGACTGTTAGAATTTACGGGC

2600 DRA3,

2642 CysGlnValProSerProGluPhePheThrGluLeuAspGlyValArgLeuHisArgPhe  
TGCCAGGTCCCATCGCCCAATTTTTCACAGAATTGGACGGGGTGCCTACATAGGTTT  
ACGGTCCAGGGTAGCGGGCTTAAAAAGTGTCTTAACCTGCCCCACGGGATGTATCCAAA

2702 AlaProProCysLysProLeuLeuArgGluGluValSerPheArgValGlyLeuHisGlu  
GCGCCCCCTGCAAGCCCTTGCTGCGGGAGGAGGTATCATTGAGAGTAGGACTCCACGAA  
CGCGGGGGGACGTTTCGGGAACGACGCCCTCCTCCATAGTAAGTCTCATCCTGAGGTGCTT

2762 TyrProValGlySerGlnLeuProCysGluProGluProAspValAlaValLeuThrSer  
TACCCGGTAGGGTCGCAATTACCTTGCGAGCCCGAACCGGACGTGGCCGTGTTGACGTCC  
ATGGGCCATCCCAGCGTTAATGGAACGCTCGGGCTTGGCCTGCACCGGCACAACCTGCAGG

2763 HGIE2, 2815 AAT2,

2822 MetLeuThrAspProSerHisIleThrAlaGluAlaAlaGlyArgArgLeuAlaArgGly  
ATGCTCACTGATCCCTCCCATATAACAGCAGAGGCGGCGGGCGAAGGTTGGCGAGGGGA  
TACGAGTGAAGTAGGGAGGTATATTGTCGTCTCCGCGGCGCGCTTCCAACCGCTCCCTT

2856 EAG1 XMA3,

2882 SerProProSerValAlaSerSerSerAlaSerGlnLeuSerAlaProSerLeuLysAla  
TCACCCCCTCTGTGGCCAGCTCCTCGGCTAGCCAGCTATCCGCTCCATCTCTCAAGGCA  
AGTGGGGGGAGACACCGGTCGAGGAGCCGATCGGTGATAGGCGAGGTAGAGAGTTCGCT

2895 BALI, 2909 NHEI,

2942 ThrCysThrAlaAsnHisAspSerProAspAlaGluLeuIleGluAlaAsnLeuLeuTrp  
ACTTGACCCGCTAACCATGACTCCCCTGATGCTGAGCTCATAGAGGCCAACCTCCTATGG  
TGAACGTGGCGATTGGTACTGAGGGGACTACGACTCGAGTATCTCCGGTTGGAGGATACC

2972 ESP1, 2975 SACI,

3002 ArgGlnGluMetGlyGlyAsnIleThrArgValGluSerGluAsnLysValValIleLeu  
AGGCAGGAGATGGGCGGCAACATCACCAGGGTTGAGTCAGAAAACAAAGTGATTCTG  
TCCGTCTCTACCCGCCGTTGTAGTGGTCCCAACTCAGTCTTTTGTTCACCACTAAGAC

3062 AspSerPheAspProLeuValAlaGluGluAspGluArgGluIleSerValProAlaGlu  
GACTCCTTCGATCCGCTTGTGGCGGAGGAGGACGAGCGGGAGATCTCCGTACCCGAGAA  
CTGAGGAAGCTAGGCGAACACCGCCTCCTGCTCGCCCTCTAGAGGCATGGGCGTCTT

3102 BGL2,

# FIGURE 11 - Page 6

3122 IleLeuArgLysSerArgArgPheAlaGlnAlaLeuProValTrpAlaArgProAspTyr  
 ATCCTGCGGAAGTCTCGGAGATTGCGCCAGGCCCTGCCCGTTTGGGCGCGGCCGGACTAT  
 TAGGACGCCTTCAGAGCCTCTAAGCGGGTCCGGGACGGGCAAACCCGCGCCGCCTGATA  
 3149 ALWN1, 3170 EAG1 XMA3,  
 3182 AsnProProLeuValGluThrTrpLysLysProAspTyrGluProProValValHisGly  
 AACCCCCCGCTAGTGGAGACGTGGAAAAAGCCCGACTACGAACCACCTGTGGTCCATGGC  
 TTGGGGGGCGATCACCTCTGCACCTTTTTCGGGCTGATGCTTGGTGGACACCAGGTACCG  
 3223 HGIE2, 3235 NCOI,  
 3242 CysProLeuProProProLysSerProProValProProProArgLysLysArgThrVal  
 TGCCCGCTTCCACCTCCAAAGTCCCCTCCTGTGCCTCCGCCTCGGAAGAAGCGGACGGTG  
 ACGGGCGAAGGTGGAGGTTTCAGGGGAGGACACGGAGGCGGAGCCTTCTTCGCCTGCCAC  
 3302 ValLeuThrGluSerThrLeuSerThrAlaLeuAlaGluLeuAlaThrArgSerPheGly  
 GTCCTCACTGAATCAACCCTATCTACTGCCTTGGCCGAGCTCGCCACCAGAAGCTTTGGC  
 CAGGAGTGACTTAGTTGGGATAGATGACGGAACCGGCTCGAGCGGTGGTCTTCGAACCG  
 3338 SACI, 3352 HIND3,  
 3362 SerSerSerThrSerGlyIleThrGlyAspAsnThrThrThrSerSerGluProAlaPro  
 AGCTCCTCAACTTCCGGCATTACGGGCGACAATACGACAACATCCTCTGAGCCCGCCCT  
 TCGAGGAGTTGAAGGCCGTAATGCCCGCTGTTATGCTGTTGTAGGAGACTCGGGCGGGGA  
 3422 SerGlyCysProProAspSerAspAlaGluSerTyrSerSerMetProProLeuGluGly  
 TCTGGCTGCCCCCGACTCCGACGCTGAGTCTTATCCTCCATGCCCCCCTGGAGGGG  
 AGACCGACGGGGGGGCTGAGGCTGCGACTCAGGATAAGGAGGTACGGGGGGGACCTCCCC  
 3443 EAM11051,  
 3482 GluProGlyAspProAspLeuSerAspGlySerTrpSerThrValSerSerGluAlaAsn  
 GAGCCTGGGGATCCGGATCTTAGCGACGGGTCATGGTCAACGGTCAGTAGTGAGGCCAAC  
 CTCGGACCCCTAGGCCTAGAATCGCTGCCAGTACCAGTTGCCAGTCATCACTCCGGTTG  
 3490 BAMHI, 3491 BSAB1, 3493 BSPE1,  
 3542 AlaGluAspValValCysCysSerMetSerTyrSerTrpThrGlyAlaLeuValThrPro  
 GCGGAGGATGTCGTGTGCTCAATGTCTTACTCTTGGACAGGCGCACTCGTCACCCCG  
 CGCTCCTACAGCACACGACGAGTTACAGAATGAGAACCTGTCCGCGTGAGCAGTGGGGC  
 3595 DRA3,  
 3602 CysAlaAlaGluGluGlnLysLeuProIleAsnAlaLeuSerAsnSerLeuLeuArgHis  
 TGCGCCGCGGAAGAACAGAACTGCCCATCAATGCACTAAGCAACTCGTTGCTACGTCAC  
 ACGCGGCGCCTTCTTGTCTTTGACGGGTAGTTACGTGATTGTTGAGCAACGATGCAGTG  
 3606 SAC2, 3617 ALWN1, 3661 PFLM1,  
 3662 HisAsnLeuValTyrSerThrThrSerArgSerAlaCysGlnArgGlnLysLysValThr  
 CACAATTTGGTGTATTCCACCACCTCACGCAGTGCTTGCCAAAGGCAGAAAGTCA  
 GTGTTAAACCACATAAGGTGGTGGAGTGCCTCACGAACGGTTTCCGTCTTCTTTCAGTGT  
 3687 DRA3,  
 PheAspArgLeuGlnValLeuAspSerHisTyrGlnAspValLeuLysGluValLysAla

3722 TTTGACAGACTGCAAGTTCTGGACAGCCATTACCGAGCTACTCAAGGAGGTTAAAGCA  
AAACTGTCTGACGTTCAAGACCTGTCGGTAATGGTCTGCATGAGTTCCTCCAATTCGT

3782 AlaAlaSerLysValLysAlaAsnLeuLeuSerValGluGluAlaCysSerLeuThrPro  
GCGGCGTCAAAAGTGAAGGCTAACTTGCTATCCGTAGAGGAAGCTTGACGCCTGACGCCC  
CGCCGCAGTTTTTCACTTCCGATTGAACGATAGGCATCTCCTTTCGAACGTGCGACTGCGGG  
3822 HIND3,

3842 ProHisSerAlaLysSerLysPheGlyTyrGlyAlaLysAspValArgCysHisAlaArg  
CCACACTCAGCCAAATCCAAGTTTGTTATGGGGCAAAGACGTCCGTTGCCATGCCAGA  
GGTGTGAGTCGGTTTAGGTTCAAACCAATACCCCGTTTTCTGCAGGCAACGGTACGGTCT  
3881.AAT2, 3896 BGLI,

3902 LysAlaValThrHisIleAsnSerValTrpLysAspLeuLeuGluAspAsnValThrPro  
AAGGCCGTAACCCACATCAACTCCGTGTGGAAAGACCTTCTGGAAGACAATGTAACACCA  
TTCCGGCATTGGGTGTAGTTGAGGCACACCTTTCTGGAAGACCTTCTGTTACATTGTGGT

3962 IleAspThrThrIleMetAlaLysAsnGluValPheCysValGlnProGluLysGlyGly  
ATAGACACTACCATCATGGCTAAGAACGAGGTTTTCTGCGTTCAGCCTGAGAAGGGGGGT  
TATCTGTGATGGTAGTACCGATTCTTGCTCCAAAAGACGCAAGTCGGACTCTTCCCCCA

4022 ArgLysProAlaArgLeuIleValPheProAspLeuGlyValArgValCysGluLysMet  
CGTAAGCCAGCTCGTCTCATCGTGTTCCCCGATCTGGGCGTGCGCGTGTGCGAAAAGATG  
GCATTCCGTCGAGCAGAGTAGCACAAAGGGGCTAGACCCGCACGCCACACGCTTTTCTAC

4082 AlaLeuTyrAspValValThrLysLeuProLeuAlaValMetGlySerSerTyrGlyPhe  
GCTTTGTACGACGTGGTTACAAAGCTCCCCTTGCCCGTGATGGGAAGCTCCTACGGATTG  
CGAAACATGCTGCACCAATGTTTCGAGGGGAACCGGCACTACCCTTCGAGGATGCCTAAG

4142 GlnTyrSerProGlyGlnArgValGluPheLeuValGlnAlaTrpLysSerLysLysThr  
CAATACTCACCAGGACAGCGGGTTGAATTCCTCGTGCAAGCGTGGAAGTCCAAGAAAACC  
GTTATGAGTGGTCTGTGCGCCCACTTAAGGAGCACGTTTCGCACCTTCAGGTTCTTTTGG  
4166 ECORI,

4202 ProMetGlyPheSerTyrAspThrArgCysPheAspSerThrValThrGluSerAspIle  
CCAATGGGGTTCTCGTATGATACCCGCTGCTTTGACTCCACAGTCACTGAGAGCGACATC  
GGTTACCCCAAGAGCATACTATGGGCGACGAAACTGAGGTGTCACTGACTCTCGTGTAG  
4235 DRD1, 4242 ALWN1,

4262 ArgThrGluGluAlaIleTyrGlnCysCysAspLeuAspProGlnAlaArgValAlaIle  
CGTACGGAGGAGGCAATCTACCAATGTTGTGACCTCGACCCCCAAGCCCGGTGGCCATC  
GCATGCCTCCTCCGTTAGATGGTTACAACACTGGAGCTGGGGGTTTCGGGCGCACCGGTAG  
4307 BGLI, 4314 BALI,

4322 LysSerLeuThrGluArgLeuTyrValGlyGlyProLeuThrAsnSerArgGlyGluAsn  
AAGTCCCTCACCAGAGAGGCTTTATGTTGGGGGCCCTCTTACCAATTCAAGGGGGGAGAAC  
TTCAGGGAGTGGCTCTCCGAAATACAACCCCGGGGAGAATGGTTAAGTTCCCCCCTCTTG  
4351 APAI,

4382 CysGlyTyrArgArgCysArgAlaSerGlyValLeuThrThrSerCysGlyAsnThrLeu  
TGCGGCTATCGCAGGTGCCGCGCAGCGGCGTACTGACAACTAGCTGTGGTAACACCCTC



498

ValProProLeuArgAlaTrpArgHisArgAlaArgSerValArgAlaArgLeuLeuAla  
5042 GTACCGCCCTTTCGAGCTTGGAGACACCGGGCCCGAGCGTCCGCGCTAGGCTTCTGGCC  
CATGGCGGGAACGCTCGAACCTCTGTGGCCCGGGCCTCGCAGGCGCGATCCGAAGACCGG  
5070 APAI, 5097 BALI,  
ArgGlyGlyArgAlaAlaIleCysGlyLysTyrLeuPheAsnTrpAlaValArgThrLys  
5102 AGAGGAGGCAGGGCTGCCATATGTGGCAAGTACCTCTTCAACTGGGCAGTAAGAACAAG  
TCTCCTCCGTCCCGACGGTATACACCGTTCATGGAGAAGTTGACCCGTCATTCTTGTTC  
5119 NDEI,  
LeuLysLeuThrProIleAlaAlaAlaGlyGlnLeuAspLeuSerGlyTrpPheThrAla  
5162 CTCAAACTCACTCCAATAGCGGCCGCTGGCCAGCTGGACTTGTCCGGCTGGTTTCACGGCT  
GAGTTTGAGTGAGGTTATCGCCGGCGACCGGTTCGACCTGAACAGGCCGACCAAGTGCCGA  
5180 NOTI, 5181 EAG1 XMA3, 5188 BALI, 5192 PVU2,  
GlyTyrSerGlyGlyAspIleTyrHisSerValSerHisAlaArgProArgTrpIleTrp  
5222 GGCTACAGCGGGGAGACATTTATCACAGCGTGTCTCATGCCCGGCCCGCTGGATCTGG  
CCGATGTCGCCCCCTCTGTAAATAGTGTGCGACAGAGTACGGGCCGGGGCGACCTAGACC  
5246 DRA3,  
PheCysLeuLeuLeuLeuAlaAlaGlyValGlyIleTyrLeuLeuProAsnArgOP  
5282 TTTTGCCTACTCTGCTTGCTGTCAGGGGTAGGCATCTACCTCCTCCCCAACCGATGAAGG  
AAAACGGATGAGGACGAACGACGTCCCCATCCGTAGATGGAGGAGGGGTTGGCTACTTCC  
5301 PSTI, 5331 HGIE2,  
5342 TTGGGGTAAACACTCCGGCCTAAAAAAAAAAAAAAAAATCTAGAACCCGAGTCGAC  
AACCCCATTTGTGAGGCCGATTTTTTTTTTTTTTTTATAGATCTTGGGCTCAGCTG  
5378 XBAI, 5390 SALI,

FIGURE 12

~~Mo~~  
SB  
std

PAS  
C

C.1 C.2

KD<sub>cr</sub>

250

-

98

-

64

-

50

-

36

-

30

-

16

-

6

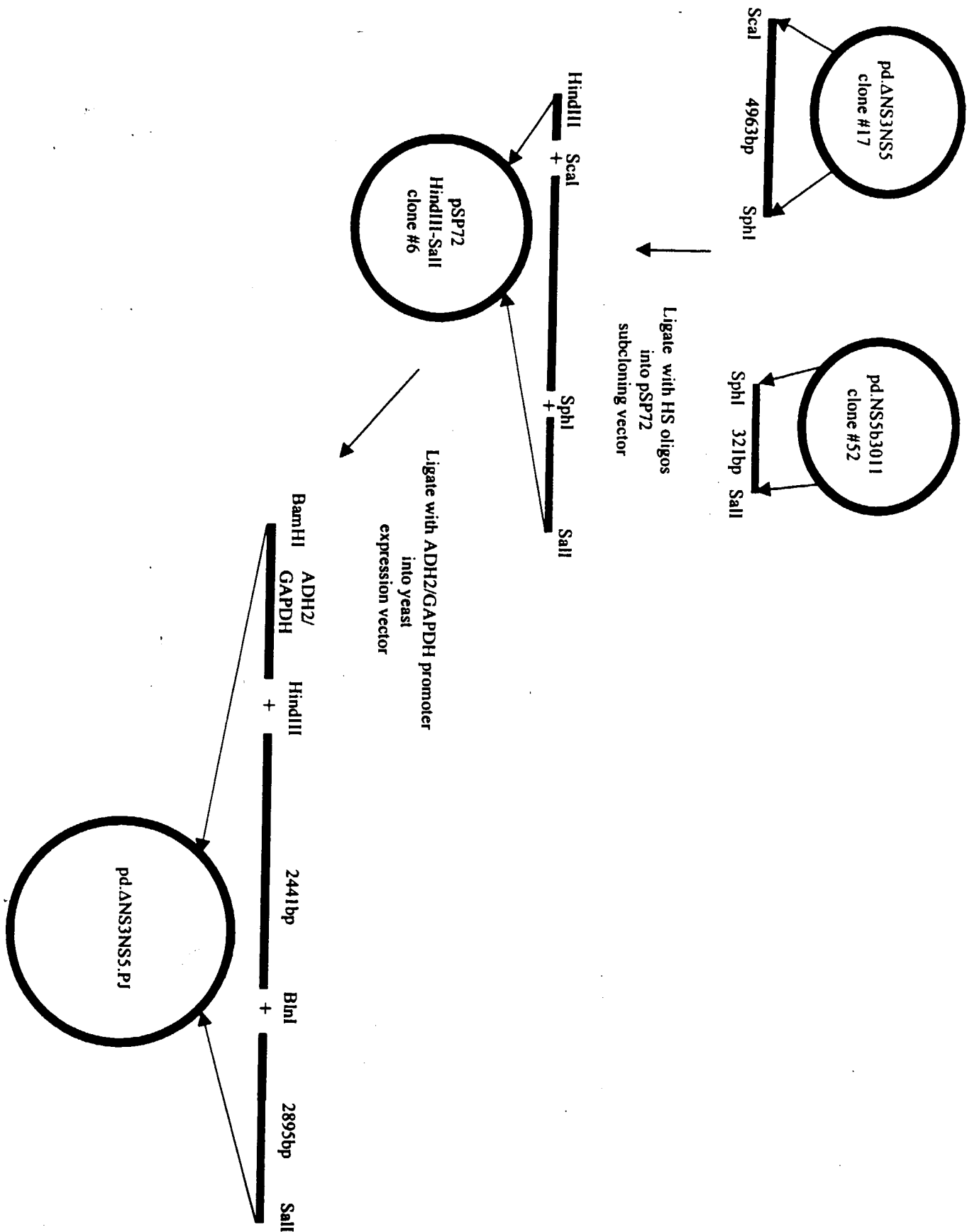
-

4

-

002275 5472260

FIGURE 13



# FIGURE 14 - Page 1

MetAlaAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsn  
 2 AGCTTACAAAACAAAATGGCTGCATATGCAGCTCAGGGCTATAAGGTGCTAGTACTCAAC  
 TCGAATGTTTTGTTTTACCGACGTATACGTCGAGTCCCGATATTCCACGATCATGAGTTG  
 ^  
 1 HIND3, 24 NDEI, 52 SCAI,  
 ProSerValAlaAlaThrLeuGlyPheGlyAlaTyrMetSerLysAlaHisGlyIleAsp  
 62 CCCTCTGTTGCTGCAACACTGGGCTTTGGTGCTTACATGTCCAAGGCTCATGGGATCGAT  
 GGGAGACAACGACGTTGTGACCCGAAACCACGAATGTACAGGTTCCGAGTACCCTAGCTA  
 ^  
 116 CLAI,  
 ProAsnIleArgThrGlyValArgThrIleThrThrGlySerProIleThrTyrSerThr  
 122 CCTAACATCAGGACCGGGGTGAGAACAATTACCACTGGCAGCCCCATCAGTACTCCACC  
 GGATTGTAGTCCTGGCCCCACTCTTGTTAATGGTGACCGTCGGGGTAGTGCATGAGGTGG  
 TyrGlyLysPheLeuAlaAspGlyGlyCysSerGlyGlyAlaTyrAspIleIleIleCys  
 182 TACGGCAAGTTCCTTGCCGACGGCGGGTGCTCGGGGGGCGCTTATGACATAATAATTTGT  
 ATGCCGTTCAAGGAACGGCTGCCGCCACGAGCCCCCGGAATACTGTATTATTAAACA  
 AspGluCysHisSerThrAspAlaThrSerIleLeuGlyIleGlyThrValLeuAspGln  
 242 GACGAGTGCCACTCCACGGATGCCACATCCATCTTGGGCATTGGCACTGTCCTTGACCAA  
 CTGCTCACGGTGAGGTGCCTACGGTGTAGGTAGAACCCGTAACCGTGACAGGAAGTGGTT  
 AlaGluThrAlaGlyAlaArgLeuValValLeuAlaThrAlaThrProProGlySerVal  
 302 GCAGAGACTGCGGGGGCGAGACTGGTTGTGCTCGCCACCGCCACCCCTCCGGGCTCCGTC  
 CGTCTCTGACGCCCCCGCTCTGACCAACACGAGCGGTGGCGGTGGGGAGGCCCGAGGCAG  
 ^  
 303 ALWN1,  
 ThrValProHisProAsnIleGluGluValAlaLeuSerThrThrGlyGluIleProPhe  
 362 ACTGTGCCCCATCCCAACATCGAGGAGTTGCTCTGTCCACCACCGGAGAGATCCCTTTT  
 TGACACGGGGTAGGGTTGTAGCTCCTCCAACGAGACAGGTGGTGGCCTCTCTAGGGAAAA  
 TyrGlyLysAlaIleProLeuGluValIleLysGlyGlyArgHisLeuIlePheCysHis  
 422 TACGGCAAGGCTATCCCCCTCGAAGTAATCAAGGGGGGAGACATCTCATCTTCTGTCAT  
 ATGCCGTTCCGATAGGGGGAGCTTCATTAGTTCCCCCCTCTGTAGAGTAGAAGACAGTA  
 SerLysLysLysCysAspGluLeuAlaAlaLysLeuValAlaLeuGlyIleAsnAlaVal  
 482 TCAAAGAAGAAGTGCAGCAACTCGCCGCAAAGCTGGTCGCATTGGGCATCAATGCCGTG  
 AGTTTCTTCTTCACGCTGCTTGAGCGGCGTTTCGACCAGCGTAACCGTAGTTACGGCAC  
 AlaTyrTyrArgGlyLeuAspValSerValIleProThrSerGlyAspValValValVal  
 542 GCCTACTACCGCGGTCTTGACGTGTCCGTCATCCCGACCAGCGCGCATGTTGTGCTCGTG  
 CGGATGATGGCGCCAGAACTGCACAGGCAGTAGGGCTGGTCGCCGCTACAACAGCAGCAC  
 ^  
 550 SAC2, 560 DRD1,  
 AlaThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsn  
 602 GCAACCGATGCCCTCATGACCGGCTATACCGGCGACTTCGACTCGGTGATAGACTGCAAT  
 CGTTGGCTACGGGAGTACTGGCCGATATGGCCGCTGAAGCTGAGCCACTATCTGACGTTA  
 ^  
 615 BSPH1,  
 ThrCysValThrGlnThrValAspPheSerLeuAspProThrPheThrIleGluThrIle

002247 " 624260

662 ACGTGTGTCAACCAGACAGTTCGATTTTCAGCCTTGACCCTACCTTACCATTGAGACAAATC  
TGCACACAGTGGGTCTGTCTCAGCTAAAGTCGGAACCTGGGATGGAAGTGGTAACTCTGTTAG

722 ThrLeuProGlnAspAlaValSerArgThrGlnArgArgGlyArgThrGlyArgGlyLys  
ACGCTCCCCCAAGATGCTGTCTCCCGCACTCAACGTGCGGGCAGGACTGGCAGGGGGGAAG  
TGCGAGGGGGTTCTACGACAGAGGGCGTGAGTTGCAGCCCCGTCTTGACCGTCCCCCTTC

782 ProGlyIleTyrArgPheValAlaProGlyGluArgProSerGlyMetPheAspSerSer  
CCAGGCATCTACAGATTTGTGGCACCGGGGAGCGCCCCCTCCGGCATGTTCCGACTCGTCC  
GGTCCGTAGATGTCTAAACACCGTGGCCCCCTCGCGGGGAGGGCCGTACAAGCTGAGCAGG

816 BGLI, 833 DRD1,

842 ValLeuCysGluCysTyrAspAlaGlyCysAlaTrpTyrGluLeuThrProAlaGluThr  
GTCCTCTGTGAGTGCTATGACGCAGGCTGTGCTTGGTATGAGCTCACGCCCGCCGAGACT  
CAGGAGACACTCACGATACTGCGTCCGACACGAACCATACTCGAGTGCGGGCGGGCTCTGA

881 SACI,

902 ThrValArgLeuArgAlaTyrMetAsnThrProGlyLeuProValCysGlnAspHisLeu  
ACAGTTAGGCTACGAGCGTACATGAACACCCCGGGGCTTCCCGTGTGCCAGGACCATCTT  
TGTCAATCCGATGCTCGCATGTACTTGTGGGGCCCCGAAGGGCACACGGTCTGAGTAA

931 SMAI XMAI,

962 GluPheTrpGluGlyValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGln  
GAATTTTGGGAGGGCGTCTTTACAGGCCTCACTCATATAGATGCCACTTCTATCCCAG  
CTTAAACCCCTCCCGCAGAAATGTCCGGAGTGAGTATATCTACGGGTGAAAGATAGGGTC

985 STUI,

1022 ThrLysGlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGlnAlaThrValCysAla  
ACAAAGCAGAGTGGGGAGAACCTTCCTTACCTGGTAGCGTACCAAGCCACCGTGTGCGCT  
TGTTTTCGTCTACCCCTCTTGGAAGGAATGGACCATCGCATGGTTCGGTGGCACACGCGA

1069 DRA3,

1082 ArgAlaGlnAlaProProProSerTrpAspGlnMetTrpLysCysLeuIleArgLeuLys  
AGGGCTCAAGCCCCCTCCCCATCGTGGGACCAGATGTGGAAGTGTTTGATTTCGCCTCAAG  
TCCCGAGTTCGGGGAGGGGGTAGCACCCCTGGTCTACACCTTCACAACTAAGCGGAGTTC

1142 ProThrLeuHisGlyProThrProLeuLeuTyrArgLeuGlyAlaValGlnAsnGluIle  
CCCACCCTCCATGGGCCAACACCCCTGCTATACAGACTGGGCGCTGTTTCAGAAATGAAATC  
GGGTGGGAGGTACCCGTTGTGGGGACGATATGTCTGACCCGCGACAAGTCTTACTTTAG

1150 NCOI,

1202 ThrLeuThrHisProValThrLysTyrIleMetThrCysMetSerAlaAspLeuGluVal  
ACCCTGACGCACCCAGTACCAAATACATCATGACATGCATGTGCGCCGACCTGGAGGTC  
TGGGACTGCGTGGGTGAGTGGTTTATGTAGTACTGTACGTACAGCCGGCTGGACCTCCAG

1230 BSPH1, 1234 DRD1, 1237 AVA3, 1245 EAG1 XMA3, 1250 DRD1,

1262 ValThrSerThrTrpValLeuValGlyGlyValLeuAlaAlaLeuAlaAlaTyrCysLeu  
GTCACGAGCACCTGGGTGCTCGTTGGCGGCGTCTGGCTGCTTTGGCCGCGTATTGCCTG  
CAGTGCTCGTGGACCCACGAGCAACCGCCGAGGACCGACGAAACCGGCGCATAACGGA

SerThrGlyCysValValIleValGlyArgValValLeuSerGlyLysProAlaIleIle  
 1322 TCAACAGGCTGCGTGGTCATAGTGGGCAGGGTCGTCTTGTCGGGAAGCCGGCAATCATA  
 AGTTGTCCGACGCACCAGTATCACCCGTCCCAGCAGAACAGGCCCTTCGGCCGTTAGTAT  
 1369 NAEI,  
 ProAspArgGluValLeuTyrArgGluPheAspGluMetGluGluCysSerGlnHisLeu  
 1382 CCTGACAGGGAAGTCCTCTACCGAGAGTTTCGATGAGATGGAAGAGTGCTCTCAGCACTTA  
 GGACTGTCCCTTCAGGAGATGGCTCTCAAGCTACTCTACCTTCTCACGAGAGTCGTGAAT  
 1385 DRD1,  
 ProTyrIleGluGlnGlyMetMetLeuAlaGluGlnPheLysGlnLysAlaLeuGlyLeu  
 1442 CCGTACATCGAGCAAGGGATGATGCTCGCCGAGCAGTTCAAGCAGAAGGCCCTCGGCCTC  
 GGCATGTAGCTCGTTCCTTACTACGAGCGGCTCGTCAAGTTCGTCTTCCGGGAGCCGGAG  
 LeuGlnThrAlaSerArgGlnAlaGluValIleAlaProAlaValGlnThrAsnTrpGln  
 1502 CTGCAGACCGCGTCCCGTCAGGCAGAGGTTATCGCCCCTGCTGTCCAGACCAACTGGCAA  
 GACGTCTGGCGCAGGGCAGTCCGTCTCCAATAGCGGGGACGACAGGTCTGGTTGACCGTT  
 1502 PSTI, 1507 TTH3I,  
 LysLeuGluThrPheTrpAlaLysHisMetTrpAsnPheIleSerGlyIleGlnTyrLeu  
 1562 AAATCGAGACCTTCTGGGCGAAGCATATGTGGAACCTTCATCAGTGGGATACAATACTTG  
 TTTGAGCTCTGGAAGACCCGCTTCGTATACACCTTGAAGTAGTCACCCTATGTTATGAAC  
 1565 XHOI, 1586 NDEI,  
 AlaGlyLeuSerThrLeuProGlyAsnProAlaIleAlaSerLeuMetAlaPheThrAla  
 1622 GCGGGCTTGTCACACGCTGCCTGGTAACCCCGCCATTGCTTCATTGATGGCTTTTACAGCT  
 CGCCCGAACAGTTGCGACGGACCATTGGGGCGGTAACGAAGTAACTACCGAAAATGTCTGA  
 1643 BSTE2, 1677 ALWN1 PVU2,  
 AlaValThrSerProLeuThrThrSerGlnThrLeuLeuPheAsnIleLeuGlyGlyTrp  
 1682 GCTGTCAACAGCCCACTAACCCTAGCCAAACCCTCCTCTTCAACATATTGGGGGGGTGG  
 CGACAGTGGTTCGGGTGATTGGTGATCGGTTTGGGAGGAGAAGTTGTATAACCCCCCACC  
 ValAlaAlaGlnLeuAlaAlaProGlyAlaAlaThrAlaPheValGlyAlaGlyLeuAla  
 1742 GTGGCTGCCCAGCTCGCCGCCCGCGTGCCGCTACTGCCTTTGTGGGCGCTGGCTTAGCT  
 CACCGACGGGTTCGAGCGGCGGGGGCCACGGCGATGACGGAAACACCCGCGACCGAATCGA  
 1794 ESP1,  
 GlyAlaAlaIleGlySerValGlyLeuGlyLysValLeuIleAspIleLeuAlaGlyTyr  
 1802 GGCGCCGCCATCGGCAGTGTTGGACTGGGGAAGGTCTCATAGACATCTTGCAGGGTAT  
 CCGCGGCGGTAGCCGTCAACCTGACCCCTTCCAGGAGTATCTGTAGGAACGTCCCAT  
 1802 KAS1 NARI,  
 GlyAlaGlyValAlaGlyAlaLeuValAlaPheLysIleMetSerGlyGluValProSer  
 1862 GGCGCGGGCGTGCGGGGAGCTCTTGTCGATTCAAGATCATGAGCGGTGAGGTCCCCCTCC  
 CCGCGCCCGCACCGCCCTCGAGAACACCGTAAGTTCTAGTACTCGCCACTCCAGGGGAGG  
 1878 SACI, 1899 BSPH1,

1922 ThrGluAspLeuValAsnLeuLeuProAlaIleLeuSerProGlyAlaLeuValValGly  
ACGGAGGACCTGGTCAATCTACTGCCGCCATCCTCTCGCCCGGAGCCCTCGTAGTCGGC  
TGCCTCCTGGACCAGTTAGATGACGGGCGGTAGGAGAGCGGGCCTCGGGAGCATCAGCCG

1928 TTH31.

ValValCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaValGlnTrp  
1982 GTGGTCTGTGCAGCAATACTGCGCCGGCACGTTGGCCCGGGCGAGGGGGCAGTGCAGTGG  
CACCAGACACGTCGTTATGACGCGGCCGTGCAACCGGGCCCGCTCCCCCGTCACGTACC

2004 NAEI, 2017 SMAI XMAI.

MetAsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValSerProThrHisTyrVal  
2042 ATGAACCGGCTGATAGCCTTCGCCTCCCGGGGAACCATGTTTCCCCACGCACACTACGTG  
TACTTGGCCGACTATCGGAAGCGGAGGGCCCCCTTGGTACAAAGGGGGTGCGTGATGCAC

2067 SMAI XMAI, 2093 DRA3,

ProGluSerAspAlaAlaAlaArgValThrAlaIleLeuSerSerLeuThrValThrGln  
2102 CCGGAGAGCGATGCAGCTGCCCGCTACTGCCATACTCAGCAGCCTCACTGTAACCCAG  
GGCCTCTCGCTACGTCGACGGGCGCAGTGACGGTATGAGTCGTCGGAGTGACATTGGGTC

2115 PVU2, 2159 ALWN1,

2162    LeuLeuArgArgLeuHisGlnTrpIleSerSerGluCysThrThrProCysSerGlySer  
CTCCTGAGGCGACTGCACCACTGGATAAGCTCGGAGTGTACCACTCCATGCTCCGGTTCC  
GAGGACTCCGCTGACGTGGTCACCTATTTCGAGCCTCACATGGTGAGGTACGAGGCCAAGG

2164 MST2, 2220 ECON1,

2222 TrpLeuArgAspIleTrpAspTrpIleCysGluValLeuSerAspPheLysThrTrpLeu  
TGGCTAAGGGACATCTGGGACTGGATATGCGAGGTGTTGAGCGACTTTAAGACCTGGCTA  
ACCGATTCCCTGTAGACCCTGACCTATACGCTCCACAACCTCGCTGAAATTCTGGACCGAT

2292 LysAlaLysLeuMetProGlnLeuProGlyIleProPheValSerCysGlnArgGlyTyr  
AAAGCTAAGCTCATGCCACAGCTGCCTGGGATCCCCCTTGTGTCTCTGCCAGCGCGGGTAT  
TTTCGATTTCGAGTACGGTGTCTGACGGACCCTAGGGGAAACACAGGACGGTTCGCGCCCAT

2285 ESP1, 2300 PVU2, 2310 BAMHI,

2342 LysGlyValTrpArgGlyAspGlyIleMetHisThrArgCysHisCysGlyAlaGluIle  
AAGGGGGTCTGGCGAGGGGACGGCATCATGCACACTCGCTGCCACTGTGGAGCTGAGATC  
TTCCCCAGACCGCTCCCTGCCGTAGTACGTGTGAGCGACGGTGACACCTCGACTCTAG

2402 ThrGlyHisValLysAsnGlyThrMetArgIleValGlyProArgThrCysArgAsnMet  
ACTGGACATGTCAAAACGGGACGATGAGGATCGTCGGTCCTAGGACCTGCAGGAACATG  
TGACCTGTACAGTTTTTGCCCTGCTACTCCTAGCAGCCAGGATCCTGGACGTCCTTGATAC

2425 BSAB1, 2441 AVR2, 2448 SSE83871, 2449 PSTI,

2462 TrpSerGlyThrPheProIleAsnAlaTyrThrThrGlyProCysThrProLeuProAla  
TGGAGTGGGACCTTCCCCATTAATGCCTACACCACGGGCCCTGTACCCCCCTTCTGCG  
ACCTCACCTGGAAGGGGTAATTACGGATGTGGTGCCCGGGACATGGGGGAAGGACGC

2480 ASE1, 2497 APAI,

ProAsnTyrThrPheAlaLeuTrpArgValSerAlaGluGluTyrValGluIleArgGln

[illegible]



2522 CCGAACTACACGTTTCGCGCTATGGAGGGTGTCTGCAGAGGAATACGTGGAGATAAGGCGAG  
GGCTTGATGTGCAAGCGCGATACCTCCCACAGACGTCTCCTTATGCACCTCTATTCCGTC

2553 PSTI,

2582 ValGlyAspPheHisTyrValThrGlyMetThrThrAspAsnLeuLysCysProCysGln  
GTGGGGGACTTCCACTACGTGACGGGTATGACTACTGACAATCTTAAATGCCCCGTGCCAG  
CACCCCCTGAAGGTGATGCACTGCCCATCTGATGACTGTTAGAATTTACGGGCGACGGTC

2594 DRA3,

2642 ValProSerProGluPhePheThrGluLeuAspGlyValArgLeuHisArgPheAlaPro  
GTCCCATCGCCCGAATTTTTCACAGAATTGGACGGGGTGCGCCTACATAGGTTTGCGCC  
CAGGGTAGCGGGCTTAAAAAGTGTCTTAACCTGCCCCACGCGGATGTATCCAAACGCGGG

2702 ProCysLysProLeuLeuArgGluGluValSerPheArgValGlyLeuHisGluTyrPro  
CCCTGCAAGCCCTTGCTGCGGGAGGAGGTATCATTGAGAGTAGGACTCCACGAATACCCG  
GGGACGTTGCGGAACGACGCCCTCCTCCATAGTAAGTCTCATCCTGAGGTGCTTATGGGC

2757 HGIE2,

2762 ValGlySerGlnLeuProCysGluProGluProAspValAlaValLeuThrSerMetLeu  
GTAGGGTCGCAATTACCTTGCGAGCCCGAACCGGACGTGGCCGTGTTGACGTCCATGCTC  
CATCCCAGCGTTAATGGAACGCTCGGGCTTGCCCTGCACCGGCACAACCTGCAGGTACGAG

2809 AAT2,

2822 ThrAspProSerHisIleThrAlaGluAlaAlaGlyArgArgLeuAlaArgGlySerPro  
ACTGATCCCTCCCATATAACAGCAGAGGCGGCCGCGGCGAAGGTTGGCGAGGGGATCACCC  
TGACTAGGGAGGGTATATTGTCGTCTCCGCCGCGCGCTTCCAACCGCTCCCCTAGTGGG

2850 EAG1 XMA3,

2882 ProSerValAlaSerSerSerAlaSerGlnLeuSerAlaProSerLeuLysAlaThrCys  
CCCTCTGTGGCCAGCTCCTCGGCTAGCCAGCTATCCGCTCCATCTCTCAAGGCAACTTGC  
GGGAGACACCGGTGAGGAGCCGATCGGTGATAGGCGAGGTAGAGAGTTCGGTTGAACG

2889 BALI, 2903 NHEI,

2942 ThrAlaAsnHisAspSerProAspAlaGluLeuIleGluAlaAsnLeuLeuTrpArgGln  
ACCGCTAACCATGACTCCCCTGATGCTGAGCTCATAGAGGCCAACCTCCTATGGAGGCAG  
TGGCGATTGGTACTGAGGGGACTACGACTCGAGTATCTCCGTTGGAGGATACCTCCGTC

2966 ESP1, 2969 SACI,

3002 GluMetGlyGlyAsnIleThrArgValGluSerGluAsnLysValValIleLeuAspSer  
GAGATGGGCGGCAACATCACCAGGGTTGAGTCAGAAAACAAAGTGGTGATTCTGGACTCC  
CTCTACCCGCCGTTGTAGTGGTCCCAACTCAGTCTTTTGTTCACCACTAAGACCTGAGG

3062 PheAspProLeuValAlaGluGluAspGluArgGluIleSerValProAlaGluIleLeu  
TTCGATCCGCTTGTGGCGGAGGAGGACGAGCGGGAGATCTCCGTACCCGCAGAAATCCTG  
AAGCTAGGCGAACACCGCCTCCTCCTGCTCGCCCTCTAGAGGCATGGGCGTCTTTAGGAC

3096 BGL2,

3122 ArgLysSerArgArgPheAlaGlnAlaLeuProValTrpAlaArgProAspTyrAsnPro  
CGGAAGTCTCGGAGATTCGCCCAGGCCCTGCCCGTTTGGGCGCGGCCGACTATAACCCC

# FIGURE 14 - Page 6

GCCTTCAGAGCCTCTAAGCGGGTCCGGGACGGGCAAACCCGCGCCGGCCTGATATTGGGG

3143 ALWN1, 3164 EAG1 XMA3,

3182 ProLeuValGluThrTrpLysLysProAspTyrGluProProValValHisGlyCysPro  
CCGCTAGTGGAGACGTGGAAAAAGCCCGACTACGAACCACTGTGGTCCATGGCTGCCCC  
GGCGATCACCTCTGCACCTTTTTCGGGCTGATGCTTGGTGGACACCAAGTACCGACGGGC

3217 HGIE2, 3229 NCOI,

3242 LeuProProProLysSerProProValProProProArgLysLysArgThrValValLeu  
CTTCCACCTCCAAAGTCCCTCCTGTGCCTCCGCCTCGGAAGAAGCGGACGGTGGTCTCTC  
GAAGGTGGAGGTTTCAGGGGAGGACACGGAGGCGGAGCCTTCTTCGCCTGCCACCAGGAG

3302 ThrGluSerThrLeuSerThrAlaLeuAlaGluLeuAlaThrArgSerPheGlySerSer  
ACTGAATCAACCCTATCTACTGCCTTGGCCGAGCTCGCCACCAGAAGCTTTGGCAGCTCC  
TGACTTAGTTGGGATAGATGACGGAACCGGCTCGAGCGGTGGTCTTCGAAACCGTCGAGG

3332 SACI, 3346 HIND3,

3362 SerThrSerGlyIleThrGlyAspAsnThrThrThrSerSerGluProAlaProSerGly  
TCAACTTCCGGCATTACGGGCGACAATACGACAACATCCTCTGAGCCCGCCCTTCTGGC  
AGTTGAAGGCCGTAATGCCCGCTGTTATGCTGTTGTAGGAGACTCGGGCGGGGAAGACCG

3422 CysProProAspSerAspAlaGluSerTyrSerSerMetProProLeuGluGlyGluPro  
TGCCCCCCCCGACTCCGACGCTGAGTCCTATTCTCCATGCCCCCCTGGAGGGGGAGCCT  
ACGGGGGGGCTGAGGCTGCGACTCAGGATAAGGAGGTACGGGGGGGACCTCCCCCTCGGA

3437 EAM11051,

3482 GlyAspProAspLeuSerAspGlySerTrpSerThrValSerSerGluAlaAsnAlaGlu  
GGGGATCCGGATCTTAGCGACGGGTCATGGTCAACGGTCAGTAGTGAGGCCAACGCGGAG  
CCCCTAGGCCTAGAATCGCTGCCAGTACCAGTTGCCAGTCATCACTCCGGTTGCGCCTC

3484 BAMHI, 3485 BSAB1, 3487 BSPE1,

3542 AspValValCysCysSerMetSerTyrSerTrpThrGlyAlaLeuValThrProCysAla  
GATGTGCTGTGCTGCTCAATGTCTTACTCTTGGACAGGCGCACTCGTCACCCCGTGCGCC  
CTACAGCACACGACGAGTTACAGAATGAGAACCTGTCCGCGTGAGCAGTGGGGCACGCGG

3589 DRA3, 3600 SAC2,

3602 AlaGluGluGlnLysLeuProIleAsnAlaLeuSerAsnSerLeuLeuArgHisHisAsn  
GCGGAAGAACAGAACTGCCCATCAATGCACTAAGCAACTCGTTGCTACGTCAACACAAT  
CGCCTTCTTGTCTTTGACGGGTAGTTACGTGATTGCTTGAGCAACGATGCAGTGGTGTTA

3611 ALWN1, 3655 PFLM1,

3662 LeuValTyrSerThrThrSerArgSerAlaCysGlnArgGlnLysLysValThrPheAsp  
TTGGTGTATTCCACCACCTCACGCAGTGCTTGCCAAAGGCAGAAGAAAGTCACATTTGAC  
AACCACATAAGGTGGTGGAGTGCCTCACGAACGGTTTCCGTCTTCTTTAGTGTAACTG

3681 DRA3,

3722 ArgLeuGlnValLeuAspSerHisTyrGlnAspValLeuLysGluValLysAlaAlaAla  
AGACTGCAAGTTCTGGACAGCCATTACCAGGACGTACTCAAGGAGGTTAAAGCAGCGCGG  
TCTGACGTTCAAGACCTGTCGGTAATGGTCTGTCATGAGTTCTCTCAATTTCTGTCGCCG

3782 SerLysValLysAlaAsnLeuLeuSerValGluGluAlaCysSerLeuThrProProHis  
TCAAAAGTGAAGGCTAACTTGCTATCCGTAGAGGAAGCTTGACGCTGACGCCCCACAC  
AGTTTTCACTTCCGATTGAACGATAGGCATCTCCTTCGAACGTCGGACTGCGGGGGTGTG  
3816 HIND3,

3842 SerAlaLysSerLysPheGlyTyrGlyAlaLysAspValArgCysHisAlaArgLysAla  
TCAGCCAAATCCAAGTTTGGTTATGGGGCAAAGACGTCCGTTGCCATGCCAGAAAGGCC  
AGTCGGTTTAGGTTCAAACCAATACCCCGTTTTCTGCAGGCAACGGTACGGTCTTTCCGG  
3875 AAT2, 3890 BGLI,

3902 ValThrHisIleAsnSerValTrpLysAspLeuLeuGluAspAsnValThrProIleAsp  
GTAACCCACATCAACTCCGTGTGGAAAGACCTTCTGGAAGACAATGTAACACCAATAGAC  
CATTGGGTGTAGTTGAGGCACACCTTTCTGGAAGACCTTCTGTTACATTGTGGTTATCTG

3962 ThrThrIleMetAlaLysAsnGluValPheCysValGlnProGluLysGlyGlyArgLys  
ACTACCATCATGGCTAAGAACGAGGTTTTCTGCGTTCAGCCTGAGAAGGGGGGTGCGTAAG  
TGATGGTAGTACCGATTCTTGCTCCAAAAGACGCAAGTCGGACTCTTCCCCCAGCATTC

4022 ProAlaArgLeuIleValPheProAspLeuGlyValArgValCysGluLysMetAlaLeu  
CCAGCTCGTCTCATCGTGTCCCCGATCTGGGCGTGCGCGTGTGCGAAAAGATGGCTTTG  
GGTCGAGCAGAGTAGCACAAAGGGGCTAGACCCGCACGCGCACACGCTTTCTACCGAAAC

4082 TyrAspValValThrLysLeuProLeuAlaValMetGlySerSerTyrGlyPheGlnTyr  
TACGACGTGGTTACAAAGCTCCCTTGCGCGTGATGGGAAGCTCCTACGGATTCCAATAC  
ATGCTGCACCAATGTTTCGAGGGGAACCGGCACTACCCTTCGAGGATGCCTAAGGTTATG

4142 SerProGlyGlnArgValGluPheLeuValGlnAlaTrpLysSerLysLysThrProMet  
TCACCAGGACAGCGGGTTGAATTCCTCGTGCAAGCGTGGAAGTCCAAGAAAACCCCAATG  
AGTGGTCTGTGCGCCCACTTAAGGAGCACGTTGCGACCTTCAGGTTCTTTTGGGGTTAC  
4160 ECORI,

4202 GlyPheSerTyrAspThrArgCysPheAspSerThrValThrGluSerAspIleArgThr  
GGGTTCTCGTATGATACCCGCTGCTTTGACTCCACAGTCACTGAGAGCGACATCCGTACG  
CCCAAGAGCATACTATGGGCGACGAACTGAGGTGTCAGTGACTCTCGCTGTAGGCATGC  
4229 DRD1, 4236 ALWN1,

4262 GluGluAlaIleTyrGlnCysCysAspLeuAspProGlnAlaArgValAlaIleLysSer  
GAGGAGGCAATCTACCAATGTTGTGACCTCGACCCCCAAGCCCGCTGGCCATCAAGTCC  
CTCCTCCGTTAGATGGTTACAACACTGGAGCTGGGGGTTGCGGCGCACCGGTAGTTACAG  
4301 BGLI, 4308 BALI,

4322 LeuThrGluArgLeuTyrValGlyGlyProLeuThrAsnSerArgGlyGluAsnCysGly  
CTCACCGAGAGGCTTTATGTTGGGGGCCCTCTTACCAATTCAAGGGGGGAGAACTGCGGC  
GAGTGGCTCTCCGAAATACAACCCCGGAGAAATGGTTAAGTTCCCCCTCTTGACCGCG  
4345 APAI,

4382 TyrArgArgCysArgAlaSerGlyValLeuThrThrSerCysGlyAsnThrLeuThrCys  
TATCGCAGGTGCCGCGGAGCGGCGTACTGACAACTAGCTGTGGTAACACCCTCACTTGC  
ATAGCGTCCACGGCGCGCTCGCCGCATGACTGTTGATCGACACCATTTGTGGGAGTGAACG

TyrIleLysAlaArgAlaAlaCysArgAlaAlaGlyLeuGlnAspCysThrMetLeuVal  
 4442 TACATCAAGGCCCGGGCAGCCTGTCGAGCCGCAGGGCTCCAGGACTGCACCATGCTCGTG  
 ATGTAGTTCCGGGGCCGTCGGACAGCTCGGCGTCCCGAGGTCCTGACGTGGTACGAGCAC  
 ^  
 4452 SMAI XMAI,  
 CysGlyAspAspLeuValValIleCysGluSerAlaGlyValGlnGluAspAlaAlaSer  
 4502 TGTGGCGACGACTTAGTTCGTTATCTGTGAAAGCGCGGGGTCCAGGAGGACGCGCGGAGC  
 ACACCGCTGCTGAATCAGCAATAGACACTTTCGCGCCCCCAGGTCCTCCTGCGCCGCTCG  
 ^ ^  
 4508 DRD1, 4511 TTH3I,  
 LeuArgAlaPheThrGluAlaMetThrArgTyrSerAlaProProGlyAspProProGln  
 4562 CTGAGAGCCTTCACGGAGGCTATGACCAGGTAATCCGCCCCCCTGGGGACCCCCACAA  
 GACTCTCGGAAGTGCCTCCGATACTGGTCCATGAGGCGGGGGGACCCCTGGGGGGTGT  
 ProGluTyrAspLeuGluLeuIleThrSerCysSerSerAsnValSerValAlaHisAsp  
 4622 CCAGAATACGACTTGGAGCTCATAACATCATGCTCCTCCAACGTGTCAGTCGCCCCACGAC  
 GGTCTTATGCTGAACCTCGAGTATTGTAGTACGAGGAGGTTGCACAGTCAGCGGGTGTCTG  
 ^  
 4637 SACI,  
 GlyAlaGlyLysArgValTyrTyrLeuThrArgAspProThrThrProLeuAlaArgAla  
 4682 GCGCTGGAAAGAGGGTCTACTACCTCACCCGTGACCCTACAACCCCCCTCGCGAGAGCT  
 CCGCGACCTTTCTCCCAGATGATGGAGTGGGCACTGGGATGTTGGGGGGAGCGCTCTCGA  
 ^  
 4731 NRUI,  
 AlaTrpGluThrAlaArgHisThrProValAsnSerTrpLeuGlyAsnIleIleMetPhe  
 4742 GCGTGGGAGACAGCAAGACACACTCCAGTCAATTCCTGGCTAGGCAACATAATCATGTTT  
 CGCACCCCTCTGTCGTTCTGTGTGAGGTCAGTTAAGGACCGATCCGTTGTATTAGTACAAA  
 AlaProThrLeuTrpAlaArgMetIleLeuMetThrHisPhePheSerValLeuIleAla  
 4802 GCCCCACACTGTGGGCGAGGATGATACTGATGACCCATTTCTTTAGCGTCCTTATAGCC  
 CGGGGGTGTGACACCCGCTCCTACTATGACTACTGGGTAAAGAAATCGCAGGAATATCGG  
 ^ ^  
 4806 PFLM1, 4807 DRA3,  
 ArgAspGlnLeuGluGlnAlaLeuAspCysGluIleTyrGlyAlaCysTyrSerIleGlu  
 4862 AGGGACAGCTTGAACAGGCCCTCGATTGCGAGATCTACGGGGCCTGCTACTCCATAGAA  
 TCCCTGGTCGAACTTGTCCGGGAGCTAACGCTCTAGATGCCCCGACGATGAGGTATCTT  
 ^  
 4893 BGL2,  
 ProLeuAspLeuProProIleIleGlnArgLeuHisGlyLeuSerAlaPheSerLeuHis  
 4922 CCACTGGATCTACCTCCAATCATTCAAAGACTCCATGGCCTCAGCGCATTTTCACTCCAC  
 GGTGACCTAGATGGAGGTTAGTAAGTTTCTGAGGTACCGGAGTCGCGTAAAAGTGAGGTG  
 ^  
 4954 NCOI,  
 SerTyrSerProGlyGluIleAsnArgValAlaAlaCysLeuArgLysLeuGlyValPro  
 4982 AGTTACTCTCCAGGTGAAATCAATAGGGTGGCCGCATGCCTCAGAAAACCTGGGGGTACCG  
 TCAATGAGAGGTCCACTTTAGTTATCCACCGGCGTACGGAGTCTTTTGAACCCCATGGC  
 ^ ^  
 5015 SPHI, 5035 KPNI,  
 ProLeuArgAlaTrpArgHisArgAlaArgSerValArgAlaArgLeuLeuAlaArgGly

5042 CCCTTGCGAGCTTGGAGACACCGGGCCCGGAGCGTCCGCGCTAGGCTTCTGGCCAGAGGA  
GGGAACGCTCGAACCTCTGTGGCCCGGGCCTCGCAGGCGCGATCCGAAGACCGGTCTCCT  
5064 APAI, 5091 BALI,  
GlyArgAlaAlaIleCysGlyLysTyrLeuPheAsnTrpAlaValArgThrLysLeuLys  
5102 GGCAGGGCTGCCATATGTGGCAAGTACCTCTTCAACTGGGCAGTAAGAACAAAGCTCAAA  
CCGTCCCGACGGTATACACCGTTCATGGAGAAGTTGACCCGTCATTCTTGTTCGAGTTT  
5113 NDEI,  
LeuThrProIleAlaAlaAlaGlyGlnLeuAspLeuSerGlyTrpPheThrAlaGlyTyr  
5162 CTCACCTCCAATAGCGGCCGCTGGCCAGCTGGACTTGTCCGGCTGGTTCACGGCTGGCTAC  
GAGTGAGGTTATCGCCGGCGACCGGTTCGACCTGAACAGGCCGACCAAGTGCCGACCGATG  
5174 NOTI, 5175 EAGI XMA3, 5182 BALI, 5186 PVU2,  
SerGlyGlyAspIleTyrHisSerValSerHisAlaArgProArgTrpIleTrpPheCys  
5222 AGCGGGGGAGACATTTATCACAGCGTGTCTCATGCCCCGGCCCCGCTGGATCTGGTTTTGC  
TCGCCCCCTCTGTAAATAGTGTGCGACAGAGTACGGGCCGGGGCGACCTAGACCAAAACG  
5240 DRA3,  
LeuLeuLeuLeuAlaAlaGlyValGlyIleTyrLeuLeuProAsnArgOP  
5282 CTACTCCTGCTTGCTGCAGGGGTAGGCATCTACCTCCTCCCCAACCGATGAATAGTCGAC  
GATGAGGACGAACGACGTCCCCATCCGTAGATGGAGGAGGGGTTGGCTACTTATCAGCTG  
5295 PSTI, 5336 SALI,

CCCTTGCGAGCTTGGAGACACCGGGCCCGGAGCGTCCGCGCTAGGCTTCTGGCCAGAGGA

FIGURE 15



SCANNED, # 14

FIGURE 16 - Page 1

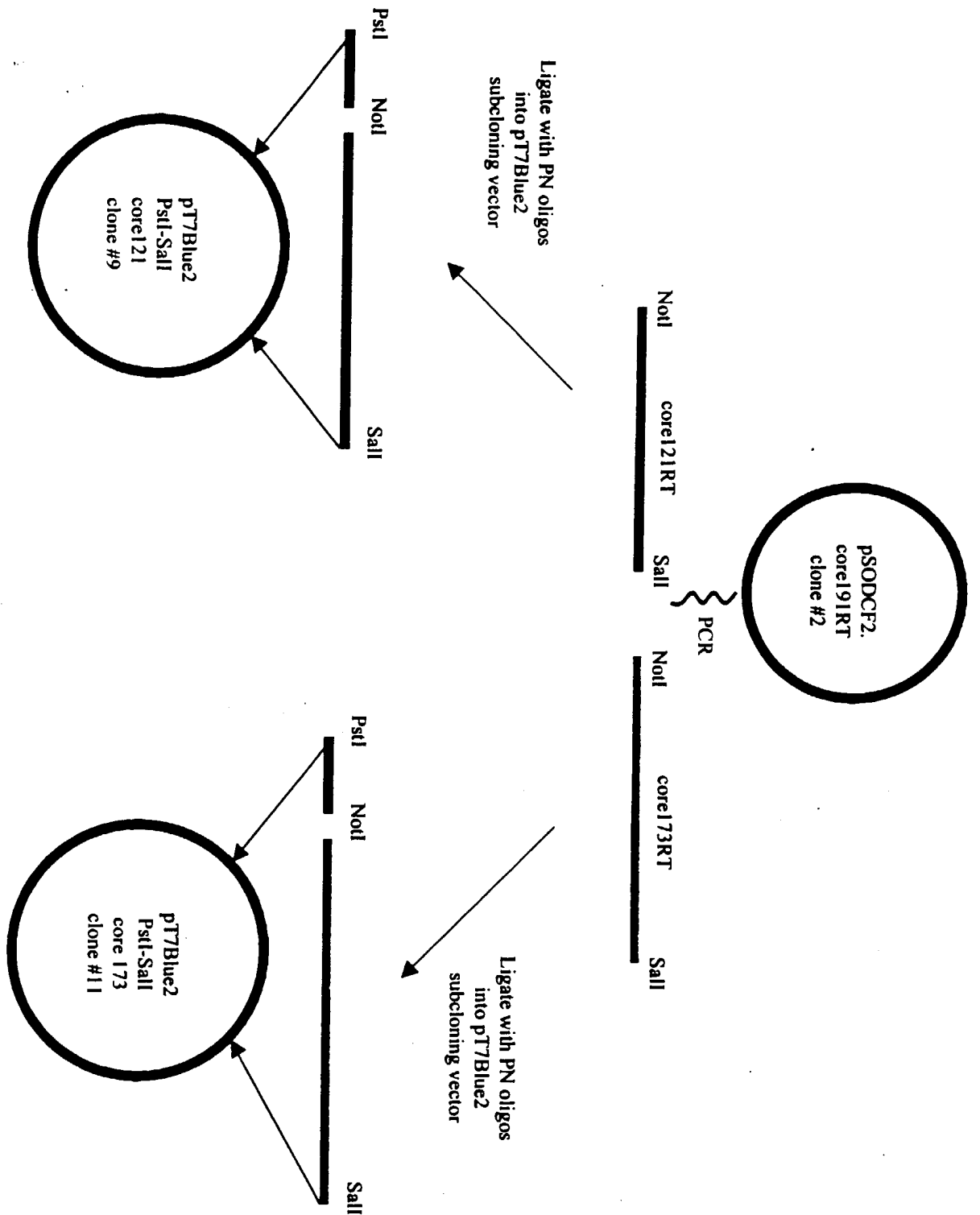
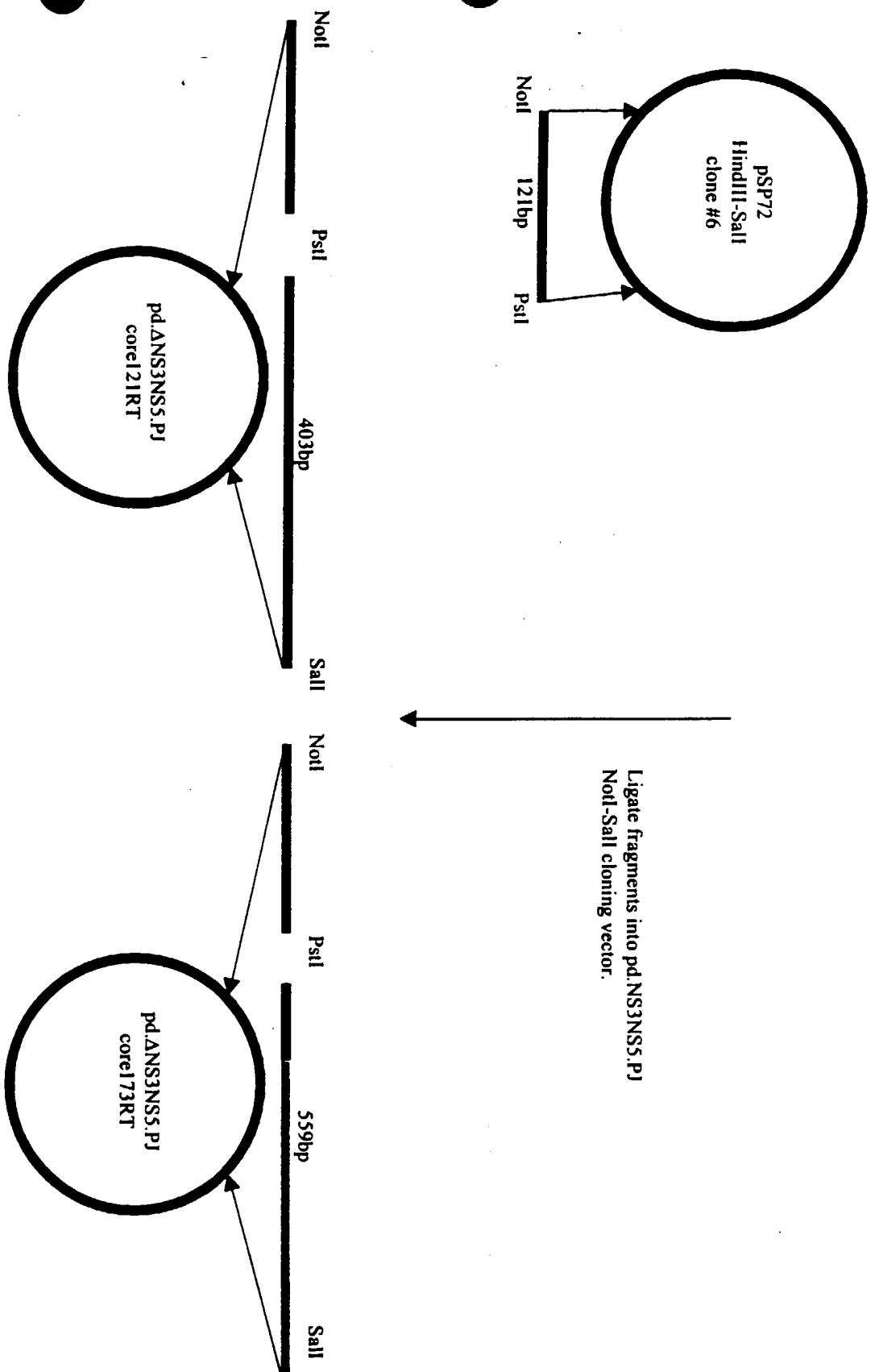


FIGURE 16 - Pa



Ligate fragments into pd.ΔNS3NS5.PJ  
NotI-SalI cloning vector.



[illegible]

MetAlaAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsn

2 AGCTTACAAAACAAAATGGCTGCATATGCAGCTCAGGGCTATAAGGTGCTAGTACTCAAC  
TCGAATGTTTTGTTTTACCGACGTATACGTCGAGTCCCGATATTCCACGATCATGAGTTG  
^ ^ ^  
1 HIND3, 24 NDEI, 52 SCAI,

ProSerValAlaAlaThrLeuGlyPheGlyAlaTyrMetSerLysAlaHisGlyIleAsp

62 CCCTCTGTTGCTGCAACACTGGGCTTTGGTGCTTACATGTCCAAGGCTCATGGGATCGAT  
GGGAGACAACGACGTTGTGACCCGAAACCACGAATGTACAGGTTCCGAGTACCCTAGCTA  
^  
116 CLAI,

ProAsnIleArgThrGlyValArgThrIleThrThrGlySerProIleThrTyrSerThr

122 CCTAACATCAGGACCGGGGTGAGAACAATTACCACTGGCAGCCCCATCACGTACTCCACC  
GGATTGTAGTCCTGGCCCCACTCTTGTTAATGGTGACCGTCGGGGTAGTGCATGAGGTGG

TyrGlyLysPheLeuAlaAspGlyGlyCysSerGlyGlyAlaTyrAspIleIleIleCys

182 TACGGCAAGTTCTTTGCCGACGGCGGGTGCTCGGGGGGCGCTTATGACATAATAATTTGT  
ATGCCGTTCAAGGAACGGCTGCCGCCACGAGCCCCCGGAATACTGTATTATTAAACA

AspGluCysHisSerThrAspAlaThrSerIleLeuGlyIleGlyThrValLeuAspGln

242 GACGAGTGCCACTCCACGGATGCCACATCCATCTTGGGCATTGGCACTGTCCTTGACCAA  
CTGCTCACGGTGAGGTGCCTACGGTGTAGGTAGAACCGTAACCGTGACAGGAAC TGTT

AlaGluThrAlaGlyAlaArgLeuValValLeuAlaThrAlaThrProProGlySerVal

302 GCAGAGACTGCGGGGGCGAGACTGGTTGTGCTCGCCACCGCCACCCCTCCGGGCTCCGTC  
CGTCTCTGACGCCCCCGCTCTGACCAACACGAGCGGTGGCGGTGGGGAGGCCCGAGGCAG  
^  
303 ALWN1,

ThrValProHisProAsnIleGluGluValAlaLeuSerThrThrGlyGluIleProPhe

362 ACTGTGCCCCATCCCAACATCGAGGAGGTTGCTCTGTCCACCACCGGAGAGATCCCTTTT  
TGACACGGGGTAGGGTTGTAGCTCCTCCAACGAGACAGGTGGTGCCCTCTCTAGGGAAAA

TyrGlyLysAlaIleProLeuGluValIleLysGlyGlyArgHisLeuIlePheCysHis

422 TACGGCAAGGCTATCCCCCTCGAAGTAATCAAGGGGGGGAGACATCTCATCTTCTGTCAT  
ATGCCGTTCCGATAGGGGGAGCTTCATTAGTTCCCCCCCCTCTGTAGAGTAGAAGACAGTA

SerLysLysLysCysAspGluLeuAlaAlaLysLeuValAlaLeuGlyIleAsnAlaVal  
 482 TCAAAGAAGAAGTGCACGAACTCGCCGCAAAGCTGGTCGCATTGGGCATCAATGCCGTG  
 AGTTTCTTCTTCACGCTGCTTGAGCGGCGTTTCGACCAGCGTAACCCGTAGTTACGGCAC  
  
 AlaTyrTyrArgGlyLeuAspValSerValIleProThrSerGlyAspValValValVal  
 542 GCCTACTACCGCGGTCTTGACGTGTCCGTATCCCGACCAGCGGCGATGTTGTCTGTCGTG  
 CGGATGATGGCGCCAGAACTGCACAGGCAGTAGGGCTGGTCGCCGCTACAACAGCAGCAC  
 550 SAC2, 560 DRD1,  
  
 AlaThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsn  
 602 GCAACCGATGCCCTCATGACCGGCTATACCGGCGACTTCGACTCGGTGATAGACTGCAAT  
 CGTTGGCTACGGGAGTACTGGCCGATATGGCCGCTGAAGCTGAGCCACTATCTGACGTTA  
 615 BSPH1,  
  
 ThrCysValThrGlnThrValAspPheSerLeuAspProThrPheThrIleGluThrIle  
 662 ACGTGTGTACCCAGACAGTCGATTTTCAGCCTTGACCCTACCTTCACCATTGAGACAATC  
 TGCACACAGTGGGTCTGTACAGCTAAAGTCGGAAGTGGGATGGAAGTGGTAACTCTGTTAG  
  
 ThrLeuProGlnAspAlaValSerArgThrGlnArgArgGlyArgThrGlyArgGlyLys  
 722 ACGCTCCCCAAGATGCTGTCTCCCGCACTCAACGTCGGGGCAGGACTGGCAGGGGGAAG  
 TGCGAGGGGGTTCTACGACAGAGGGCGTGAGTTGCAGCCCCGTCTGACCGTCCCCCTTC  
  
 ProGlyIleTyrArgPheValAlaProGlyGluArgProSerGlyMetPheAspSerSer  
 782 CCAGGCATCTACAGATTTGTGGCACCGGGGGAGCGCCCTCCGGCATGTTTCGACTCGTCC  
 GGTCCGTAGATGTCTAAACACCGTGGCCCCCTCGCGGGGAGGCCGTACAAGCTGAGCAGG  
 816 BGLI, 833 DRD1,  
  
 ValLeuCysGluCysTyrAspAlaGlyCysAlaTrpTyrGluLeuThrProAlaGluThr  
 842 GTCCTCTGTGAGTGCTATGACGCAGGCTGTGCTTGGTATGAGCTCACGCCCCGCGAGACT  
 CAGGAGACACTCACGATACTGCGTCCGACACGAACCATACTCGAGTGCGGGCGGCTCTGA  
 881 SACI,  
  
 ThrValArgLeuArgAlaTyrMetAsnThrProGlyLeuProValCysGlnAspHisLeu  
 902 ACAGTTAGGCTACGAGCGTACATGAACACCCCGGGGCTTCCCGTGTGCCAGGACCATCTT  
 TGTCAATCCGATGCTCGCATGTACTTGTGGGGCCCCGAAGGGCACACGGTCTGGTAGAA  
 931 SMAI XMAI,  
  
 GluPheTrpGluGlyValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGln  
 962 GAATTTTGGGAGGGCGTCTTTACAGGCCTCACTCATATAGATGCCCACCTTTCTATCCAG  
 CTTAAAACCCTCCCGCAGAAATGTCCGGAGTGAGTATATCTACGGGTGAAAGATAGGGTC  
 985 STUI,  
  
 ThrLysGlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGlnAlaThrValCysAla  
 1022 ACAAGCAGAGTGGGGAGAACCTTCCTTACCTGGTAGCGTACCAAGCCACCGTGTGCGCT  
 TGTTTCGTCTACCCCTCTTGGAAGGAATGGACCATCGCATGGTTCGGTGGCACACGCGA  
 1069 DRA3,  
  
 ArgAlaGlnAlaProProProSerTrpAspGlnMetTrpLysCysLeuIleArgLeuLys  
 1082 AGGGCTCAAGCCCTCCCCCATCGTGGGACCAGATGTGGAAGTGTGATTGCGCTCAAG

TCCCGAGTTCGGGGAGGGGGTAGCACCTGGTCTACACCTTCACAACTAAGCGGAGTTC

ProThrLeuHisGlyProThrProLeuLeuTyrArgLeuGlyAlaValGlnAsnGluIle  
1142 CCCACCCTCCATGGGCCAACCCCTGCTATACAGACTGGGCGCTGTTCAGAATGAAATC  
GGGTGGGAGGTACCCGGTTGTGGGGACGATATGTCTGACCCGCGACAAGTCTTACTTTAG

1150 NCOI,

ThrLeuThrHisProValThrLysTyrIleMetThrCysMetSerAlaAspLeuGluVal  
1202 ACCCTGACGCACCCAGTCACCAAATACATCATGACATGCATGTCGGCCGACCTGGAGGTC  
TGGGACTGCGTGGGTCACTGGTTTATGTAGTACTGTACGTACAGCCGGCTGGACCTCCAG

1230 BSPH1, 1234 DRD1, 1237 AVA3, 1245 EAG1 XMA3, 1250 DRD1,

ValThrSerThrTrpValLeuValGlyGlyValLeuAlaAlaLeuAlaAlaTyrCysLeu  
1262 GTCACGAGCACCTGGGTGCTCGTTGGCGGCGTCCTGGCTGCTTTGGCCGCGTATTGCCTG  
CAGTGCTCGTGGACCCACGAGCAACCGCCGCAGGACCGACGAAACCGGCGCATAACGGAC

1322 SerThrGlyCysValValIleValGlyArgValValLeuSerGlyLysProAlaIleIle  
TCAACAGGCTGCGTGGTCATAGTGGGCAGGGTCGTCTTGTCGGGAAGCCGGCAATCATA  
AGTTGTCCGACGCACCAGTATCACCCGTCCCAGCAGAACAGGCCCTTCGGCCGTTAGTAT

1369 NAEI,

ProAspArgGluValLeuTyrArgGluPheAspGluMetGluGluCysSerGlnHisLeu  
1382 CCTGACAGGGAAGTCCTCTACCGAGAGTTCGATGAGATGGAAGAGTGCTCTCAGCACTTA  
GGACTGTCCCTTCAGGAGATGGCTCTCAAGCTACTCTACCTTCTCACGAGAGTCGTGAAT

1385 DRD1,

ProTyrIleGluGlnGlyMetMetLeuAlaGluGlnPheLysGlnLysAlaLeuGlyLeu  
1442 CCGTACATCGAGCAAGGGATGATGCTCGCCGAGCAGTTCAAGCAGAAGGCCCTCGGCCTC  
GGCATGTAGCTCGTTCCTACTACGAGCGGCTCGTCAAGTTCGTCTTCCGGGAGCCGGAG

1502 LeuGlnThrAlaSerArgGlnAlaGluValIleAlaProAlaValGlnThrAsnTrpGln  
CTGCAGACCGCGTCCCGTCAGGCAGAGGTTATCGCCCCTGCTGTCCAGACCAACTGGCAA  
GACGTCTGGCGCAGGGCAGTCCGTCTCCAATAGCGGGGACGACAGGTCTGGTTGACCGTT

1502 PSTI, 1507 TTH3I,

1562 LysLeuGluThrPheTrpAlaLysHisMetTrpAsnPheIleSerGlyIleGlnTyrLeu  
AAACTCGAGACCTTCTGGGCGAAGCATATGTGGAAGTTCATCAGTGGGATACAATACTTG  
TTTGAGCTCTGGAAGACCCGCTTCGTATACACCTTGAAGTAGTCACCCTATGTTATGAAC

1565 XHOI, 1586 NDEI,

AlaGlyLeuSerThrLeuProGlyAsnProAlaIleAlaSerLeuMetAlaPheThrAla  
1622 GCGGGCTTGTC AACGCTGCCTGGTAACCCGCCATTGCTTCATTGATGGCTTTTACAGCT  
CGCCCGAACAGTTGCGACGGACCATTGGGGCGGTAACGAAGTAACTACCGAAAATGTCTGA

1643 BSTE2, 1677 ALWN1 PVU2,

AlaValThrSerProLeuThrThrSerGlnThrLeuLeuPheAsnIleLeuGlyGlyTrp  
1682 GCTGTCACCAGCCCACTAACCCTAGCCAAACCCTCCTCTTCAACATATTGGGGGGGTGG  
CGACAGTGGTCGGGTGATTGGTGATCGGTTTGGGAGGAGAAGTTGTATAACCCCCCACC

ValAlaAlaGlnLeuAlaAlaProGlyAlaAlaThrAlaPheValGlyAlaGlyLeuAla  
1742 GTGGCTGCCCAGCTCGCCGCCCCCGGTGCCGCTACTGCCTTTGTGGGCGCTGGCTTAGCT  
CACCGACGGGTTCGAGCGCGGGGGCCACGGCGATGACGGAAACACCCGCGACCGAATCGA  
1794 ESP1,  
GlyAlaAlaIleGlySerValGlyLeuGlyLysValLeuIleAspIleLeuAlaGlyTyr  
1802 GGCGCCGCCATCGGCAGTGTGGACTGGGGAAGGTCCTCATAGACATCCTTGCAGGGTAT  
CCGCGCGCGGTAGCCGTCACAACCTGACCCCTTCCAGGAGTATCTGTAGGAACGTCCCAT  
1802 KAS1 NARI,  
GlyAlaGlyValAlaGlyAlaLeuValAlaPheLysIleMetSerGlyGluValProSer  
1862 GGCGCGGGCGTGGCGGGAGCTCTTGTGGCATTCAAGATCATGAGCGGTGAGGTCCCTCC  
CCGCGCCCGCACCGCCCTCGAGAACACCGTAAGTTCTAGTACTCGCCACTCCAGGGGAGG  
1878 SACI, 1899 BSPH1,  
ThrGluAspLeuValAsnLeuLeuProAlaIleLeuSerProGlyAlaLeuValValGly  
1922 ACGGAGGACCTGGTCAATCTACTGCCCCGCCATCCTCTCGCCCGGAGCCCTCGTAGTCGGC  
TGCTCTCTGGACCAGTTAGATGACGGGCGGTAGGAGAGCGGGCCTCGGGAGCATCAGCCG  
1928 TTH3I,  
ValValCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaValGlnTrp  
1982 GTGGTCTGTGCAGCAATACTGCGCCGGCACGTTGGCCCCGGGCGAGGGGGCAGTGCAGTGG  
CACCAGACACGTCGTTATGACGCGGGCCGTGCAACCGGGCCCGCTCCCCCGTCACGTCACC  
2004 NAEI, 2017 SMAI XMAI,  
MetAsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValSerProThrHisTyrVal  
2042 ATGAACCGGCTGATAGCCTTCGCCTCCCGGGGGAACCATGTTTCCCCACGCACTACGTG  
TACTTGGCCGACTATCGGAAGCGGAGGGCCCCCTTGGTACAAAGGGGGTGCGTGTATGCAC  
2067 SMAI XMAI, 2093 DRA3,  
ProGluSerAspAlaAlaAlaArgValThrAlaIleLeuSerSerLeuThrValThrGln  
2102 CCGGAGAGCGATGCAGCTGCCCCGCGTCACTGCCATACTCAGCAGCCTCACTGTAACCCAG  
GGCCTCTCGCTACGTTCGACGGGCGCAGTGACGGTATGAGTCGTGCGAGTGACATTGGGTC  
2115 PVU2, 2159 ALWN1,  
LeuLeuArgArgLeuHisGlnTrpIleSerSerGluCysThrThrProCysSerGlySer  
2162 CTCCTGAGGCGACTGCACCACTGGATAAGCTCGGAGTGTACCACTCCATGCTCCGGTTCC  
GAGGACTCCGCTGACGTGGTTCACCTATTTCGAGCCTCACATGGTGAGGTACGAGGCCAAGG  
2164 MST2, 2220 ECON1,  
TrpLeuArgAspIleTrpAspTrpIleCysGluValLeuSerAspPheLysThrTrpLeu  
2222 TGGCTAAGGGACATCTGGGACTGGATATGCGAGGTGTTGAGCGACTTTAAGACCTGGCTA  
ACCGATTCCCTGTAGACCCTGACCTATACGCTCCACAACCTCGCTGAAATTCTGGACCGAT  
LysAlaLysLeuMetProGlnLeuProGlyIleProPheValSerCysGlnArgGlyTyr  
2282 AAAGCTAAGCTCATGCCACAGCTGCCTGGGATCCCCTTTGTGTCTTGCCAGCGCGGGTAT  
TTTCGATTTCGAGTACGGTTCGACGGACCCTAGGGGAAACACAGGACGGTCGCGCCCAT  
2285 ESP1, 2300 PVU2, 2310 BAMHI,

| Table 1. Demographic characteristics of the study population |             |
|--------------------------------------------------------------|-------------|
| Age (years)                                                  | 65.0 ± 10.0 |
| Gender                                                       |             |
| Male                                                         | 50 (50.0%)  |
| Female                                                       | 50 (50.0%)  |
| Education (years)                                            | 12.0 ± 2.0  |
| Marital status                                               |             |
| Married                                                      | 40 (80.0%)  |
| Single                                                       | 10 (20.0%)  |
| Occupation                                                   |             |
| Retired                                                      | 30 (60.0%)  |
| Unemployed                                                   | 20 (40.0%)  |
| Income (USD/month)                                           | 1,200 ± 300 |
| Health status                                                |             |
| Good                                                         | 30 (60.0%)  |
| Poor                                                         | 20 (40.0%)  |
| Comorbidities                                                |             |
| Hypertension                                                 | 15 (30.0%)  |
| Diabetes                                                     | 10 (20.0%)  |
| Cholesterol                                                  | 12 (24.0%)  |
| Smoking status                                               |             |
| Smoker                                                       | 10 (20.0%)  |
| Non-smoker                                                   | 40 (80.0%)  |
| Alcohol consumption                                          |             |
| Regular                                                      | 5 (10.0%)   |
| Occasional                                                   | 15 (30.0%)  |
| Never                                                        | 30 (60.0%)  |

2942 ThrAlaAsnHisAspSerProAspAlaGluLeuIleGluAlaAsnLeuLeuTrpArgGln  
 ACCGCTAACCATGACTCCCCTGATGCTGAGCTCATAGAGGCCAACCTCCTATGGAGGCAG  
 TGGCGATTGGTACTGAGGGGACTACGACTCGAGTATCTCCGTTGGAGGATACCTCCGTC  
 ^ ^  
 2966 ESP1, 2969 SACI,  
 GluMetGlyGlyAsnIleThrArgValGluSerGluAsnLysValValIleLeuAspSer  
 3002 GAGATGGGCGGCAACATCACCAGGGTTGAGTCAGAAAACAAAGTGTTGATTCTGGACTCC  
 CTCTACCCGCCGTTGTAGTGGTCCCAACTCAGTCTTTTGTTCACCACTAAGACCTGAGG  
 PheAspProLeuValAlaGluGluAspGluArgGluIleSerValProAlaGluIleLeu  
 3062 TTCGATCCGCTTGTGGCGGAGGAGGACGAGCGGGAGATCTCCGTACCCGCAGAAATCCTG  
 AAGCTAGGCGAACACCGCCTCCTCCTGCTCGCCCTCTAGAGGCATGGGCGTCTTTAGGAC  
 ^  
 3096 BGL2,  
 ArgLysSerArgArgPheAlaGlnAlaLeuProValTrpAlaArgProAspTyrAsnPro  
 3122 CGGAAGTCTCGGAGATTCGCCCAGGCCCTGCCCGTTTGGGCGCGGCCGGACTATAACCCC  
 GCCTTCAGAGCCTCTAAGCGGGTCCCGGACGGGCAAACCCGCGCCGGCCTGATATTGGGG  
 ^ ^  
 3143 ALWN1, 3164 EAG1 XMA3,  
 ProLeuValGluThrTrpLysLysProAspTyrGluProProValValHisGlyCysPro  
 3182 CCGCTAGTGGAGACGTGGAAAAAGCCCGACTACGAACCACTGTGGTCCATGGCTGCCCCG  
 GGCGATCACCTCTGCACCTTTTTCGGGCTGATGCTTGGTGGACACCAGGTACCGACGGG  
 ^ ^  
 3217 HGIE2, 3229 NCOI,  
 LeuProProProLysSerProProValProProProArgLysLysArgThrValValLeu  
 3242 CTTCACCTCCAAAGTCCCCTCCTGTGCCTCCGCCTCGGAAGAAGCGGACGGTGGTCCCTC  
 GAAGGTGGAGGTTTCAGGGGAGGACACGGAGGCGGAGCCTTCTTCGCCTGCCACCAGGAG  
 ThrGluSerThrLeuSerThrAlaLeuAlaGluLeuAlaThrArgSerPheGlySerSer  
 3302 ACTGAATCAACCCTATCTACTGCCTTGGCCGAGCTCGCCACCAGAAGCTTTGGCAGCTCC  
 TGACTTAGTTGGGATAGATGACGGAACCGGCTCGAGCGGTGGTCTTCGAAACCGTCGAGG  
 ^ ^  
 3332 SACI, 3346 HIND3,  
 SerThrSerGlyIleThrGlyAspAsnThrThrThrSerSerGluProAlaProSerGly  
 3362 TCAACTCCGGCATTACGGGCGACAATACGACAACATCCTCTGAGCCCCGCCCTTCTGGC  
 AGTTGAAGGCCGTAATGCCCCGCTGTTATGCTGTTGTAGGAGACTCGGGCGGGGAAGACCG  
 CysProProAspSerAspAlaGluSerTyrSerSerMetProProLeuGluGlyGluPro  
 3422 TGCCCCCCCCGACTCCGACGCTGAGTCCTATTCTCCATGCCCCCCTGGAGGGGGAGCCT  
 ACGGGGGGGCTGAGGCTGCGACTCAGGATAAGGAGGTACGGGGGGGACCTCCCCCTCGGA  
 ^  
 3437 EAM11051,  
 GlyAspProAspLeuSerAspGlySerTrpSerThrValSerSerGluAlaAsnAlaGlu  
 3482 GGGGATCCGGATCTTAGCGACGGGTCATGGTCAACGGTCAGTAGTGAGGCCAACGCGGAG  
 CCCCTAGGCCTAGAATCGCTGCCAGTACCAGTTGCCAGTCATCACTCCGTTGCGCCTC  
 ^ ^ ^  
 3484 BAMHI, 3485 BSAB1, 3487 BSPE1,  
 AspValValCysCysSerMetSerTyrSerTrpThrGlyAlaLeuValThrProCysAla  
 3542 GATGTGCTGTGCTGCTCAATGTCTTACTCTTGGACAGGCGCACTCGTCACCCCGTGGCGC  
 CTACAGCACACGACGAGTTACAGAATGAGAACCTGTCCGCGTGAGCAGTGGGGCACGCGG

3589 DRA3, 3600 SAC2,

3602 AlaGluGluGlnLysLeuProIleAsnAlaLeuSerAsnSerLeuLeuArgHisHisAsn  
GCGGAAGAACAGAACTGCCCATCAATGCACTAAGCAACTCGTTGCTACGTCACCACAAT  
CGCCTTCTTGTCTTTGACGGGTAGTTACGTGATTTCGTTGAGCAACGATGCAGTGGTGTTA

3611 ALWN1, 3655 PFLM1,

3662      LeuValTyrSerThrThrSerArgSerAlaCysGlnArgGlnLysLysValThrPheAsp  
TTGGTGTATTCCACCACCTCACGCAGTGCTTGCCAAAGGCAGAAGAAAGTCACATTTGAC  
AACCACATAAGGTGGTGGAGTGCCTCACGAACGGTTTCCGCTCTTCTTTCAGTGTAAACTG

3681. DRA3,

3722 ArgLeuGlnValLeuAspSerHisTyrGlnAspValLeuLysGluValLysAlaAlaAla  
AGACTGCAAGTTCTGGACAGCCATTACCAGGACGTACTCAAGGAGGTTAAAGCAGCGGCG  
TCTGACGTTCAAGACCTGTCTGGTAATGGTCTCTGCATGAGTTCCTCCAATTTCGTCGCCGC

3782 SerLysValLysAlaAsnLeuLeuSerValGluGluAlaCysSerLeuThrProProHis  
TCAAAAGTGAAGGCTAACTTGCTATCCGTAGAGGAAGCTTGCAGCCTGACGCCCCACAC  
AGTTTTCACTTCCGATTGAACGATAGGCATCTCCTTGAACGTCGGACTGCGGGGGTGTG

3816 HIND3,

3842 SerAlaLysSerLysPheGlyTyrGlyAlaLysAspValArgCysHisAlaArgLysAla  
TCAGCCAAATCCAAGTTTGGTTATGGGGCAAAGACGTCCGTTGCCATGCCAGAAAGGCC  
AGTCGGTTTAGGTTCAAACCAATACCCCGTTTCTGCAGGCAACGGTACGGTCTTCCGG

3875 AAT2, 3890 BGLI,

ValThrHisIleAsnSerValTrpLysAspLeuLeuGluAspAsnValThrProIleAsp  
3902 GTAACCCACATCAACTCCGTGTGGAAGACCTTCTGGAAGACAATGTAACACCAATAGAC  
CATTGGGTGTAGTTGAGGCACACCTTTCTGGAAGACCTTCTGTTACATTGTGGTTATCTG

3962 ThrThrIleMetAlaLysAsnGluValPheCysValGlnProGluLysGlyGlyArgLys  
ACTACCATCATGGCTAAGAACGAGGTTTTCTGCGTTTCAGCCTGAGAAGGGGGGTCGTAAG  
TGATGGTAGTACCGATTCTTGCTCCAAAAGACGCAAGTCGGACTCTTCCCCCAGCATTC

ProAlaArgLeuIleValPheProAspLeuGlyValArgValCysGluLysMetAlaLeu  
4022 CCAGCTCGTCTCATCGTGTTCCCGATCTGGGCGTGCGCGTGTGCGAAAAGATGGCTTTG  
GGTCGAGCAGAGTAGCACAAAGGGGCTAGACCCGCACGCGCACACGCTTTTCTACCGAAAC

4082 TyrAspValValThrLysLeuProLeuAlaValMetGlySerSerTyrGlyPheGlnTyr  
TACGACGTGGTTACAAAGCTCCCCCTGGCCGTGATGGGAAGCTCCTACGGATTCCAATAC  
ATGCTGCACCAATGTTTCGAGGGGAACCGGCACTACCCTTCGAGGATGCCTAAGGTTATG

4142 SerProGlyGlnArgValGluPheLeuValGlnAlaTrpLysSerLysLysThrProMet  
TCACCAGGACAGCGGGTTGAATTCCTCGTGAAGCGTGGAAGTCCAAGAAACCCCAATC  
AGTGGTCTGTGCCCCAACTTAAGGAGCACGTTGCGACCTTCAGGTTCTTTTGGGGTTAC

4160 ECORI,

4202 GlyPheSerTyrAspThrArgCysPheAspSerThrValThrGluSerAspIleArgThr  
GGGTTCTCGTATGATACCCGCTGCTTTGACTCCACAGTCACTGAGAGCGACATCCGTACC  
CCCAAGAGCATACTATGGGCGACGAACTGAGGTGTCAGTGACTCTCGCTGTAGGCATG

[illegible]

ArgAspGlnLeuGluGlnAlaLeuAspCysGluIleTyrGlyAlaCysTyrSerIleGlu



4862 AGGGACCAGCTTGAACAGGCCCTCGATTGCGAGATCTACGGGGCCTGCTACTCCATAGAA  
TCCCTGGTCGAACTTGTCGGGAGCTAACGCTCTAGATGCCCGGACGATGAGGTATCTT

4893 BGL2,

ProLeuAspLeuProProIleIleGlnArgLeuHisGlyLeuSerAlaPheSerLeuHis  
4922 CCACTGGATCTACCTCCAATCATTCAAAGACTCCATGGCCTCAGCGCATTTCCTACTCCAC  
GGTGACCTAGATGGAGGTTAGTAAGTTTCTGAGGTACCGGAGTCGCGTAAAAGTGAGGTG

4954 NCOI.

4982 SerTyrSerProGlyGluIleAsnArgValAlaAlaCysLeuArgLysLeuGlyValPro  
AGTTACTCTCCAGGTGAAATCAATAGGGTGGCCGCATGCCTCAGAAAACCTGGGGTACCG  
TCAATGAGAGGTCCACTTTAGTTATCCACCGGCGTACGGAGTCTTTTGAACCCCATGGC

5015 SPHI, 5035 KPNI,

ProLeuArgAlaTrpArgHisArgAlaArgSerValArgAlaArgLeuLeuAlaArgGly  
5042 CCCTTGCGAGCTTGAGACACCGGGCCGAGCGTCCGCGCTAGGCTTCTGGCCAGAGGA  
GGGAACGCTCGAACCTCTGTGGCCCGGGCCTCGCAGGCGCGATCCGAAGACCGGTCTCCT

5064 APAI, 5091 BALI,

5102 GlyArgAlaAlaIleCysGlyLysTyrLeuPheAsnTrpAlaValArgThrLysLeuLys  
GGCAGGGCTGCCATATGTGGCAAGTACCTCTTCAACTGGGCGTAGAAGAAACAAAGCTCAA  
CCGTCCCGACGGTATACACCGTTCATGGAGAAGTTGACCCGTCATTCTTGTTTCGAGTTT

5113 NDEI,

LeuThrProIleAlaAlaAlaGlyGlnLeuAspLeuSerGlyTrpPheThrAlaGlyTyr  
5162 CTCACTCCAATAGCGGCCGCTGGCCAGCTGGACTTGTCGGGCTGGTTACGGCTGGCTAC  
GAGTGAGGTTATCGCCGGCAGCCGGTGCACCTGAACAGGCCGACCAAGTGCCGACCGATG

5174 NOTI, 5175 EAG1 XMA3, 5182 BALI, 5186 PVU2,

5222 SerGlyGlyAspIleTyrHisSerValSerHisAlaArgProArgTrpIleTrpPheCys  
AGCGGGGGAGACATTTATCACAGCGTGTCTCATGCCGGCCCCGCTGGATCTGGTTTTGC  
TCGCCCCCTCTGTAAATAGTGTGCGACAGAGTACGGGCCGGGGCGACCTAGACCAAACG

5240 DRA3,

5282 LeuLeuLeuLeuAlaAlaGlyValGlyIleTyrLeuLeuProAsnArgMetSerThrAsn  
CTACTCCTGCTTGCTGCAGGGGTAGGCATCTACCTCCTCCCCAACCGAATGAGCACGAAT  
GATGAGGACGAACGACGTCCCCATCCGTAGATGGAGGAGGGGTGGCTTACTCGTGCTTA

5295 PSTI,

ProLysProGlnArgLysThrLysArgAsnThrAsnArgArgProGlnAspValLysPhe  
5342 CCTAAACCTCAAAGAAAGACCAAACGTAACACCAACCGCGCGGCCGAGGACGTCAAGTTC  
GGATTTGGAGTTTCTTTCTGGTTTGCATTGTGGTTGGCCGCCGGCGTCTCTGCAGTTCAAG

5380 NOT1, 5381 EAG1 XMA3, 5390 AAT2, 5401 SMAI XMAI,

ProGlyGlyGlyGlnIleValGlyGlyValTyrLeuLeuProArgArgGlyProArgLeu  
5402 CCGGGTGGCGGT CAGATCGTTGGTGGAGTTTACTTGTTGCCGCGCAGGGGCCCTAGATTG  
GGCCACCGCCAGTCTAGCAACCACCTCAAATGAACAACGGCGCGTCCCCGGGATCTAAC

# FIGURE 17 - Page 10

5449 APAI,

5462 GlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGlyArgArgGlnPro  
GGTGTGCGCGCGACGAGAAAGACTTCCGAGCGGTGCGAACCTCGAGGTAGACGTCAGCCT  
CCACACGCGCGCTGCTCTTTCTGAAGGCTCGCCAGCGTTGGAGCTCCATCTGCAGTCGGA

5467 BSSH2, 5478 XMNI, 5502 XHOI, 5511 AAT2,

5522 IleProLysAlaArgArgProGluGlyArgThrTrpAlaGlnProGlyTyrProTrpPro  
ATCCCCAAGGCTCGTCGGCCCCGAGGGCAGGACCTGGGCTCAGCCCGGGTACCCTTGGCCC  
TAGGGGTTCGAGCAGCCGGGCTCCCGTCCTGGACCCGAGTCGGGCCCATGGGAACCGGG

5548 ALWN1, 5558 ESP1, 5564 SMAI XMAI, 5568 KPNI,

5582 LeuTyrGlyAsnGluGlyCysGlyTrpAlaGlyTrpLeuLeuSerProArgGlySerArg  
CTCTATGGCAATGAGGGCTGCGGGTGGGCGGGATGGCTCCTGTCTCCCCGTGGCTCTCGG  
GAGATACCGTTACTCCCGACGCCACCCGCCCTACCGAGGACAGAGGGGCACCGAGAGCC

5642 ProSerTrpGlyProThrAspProArgArgArgSerArgAsnLeuGlyLysOC AM  
CCTAGCTGGGGCCCCACAGACCCCCGGCGTAGGTGCGCAATTTGGGTAAGTAATAGTCG  
GGATCGACCCCGGGGTGTCTGGGGGCCGATCCAGCGCGTTAAACCCATTTCATTATCAGC

5650 APAI, 5698 SALI,

5702 AC  
TG

5449 APAI, 5462 GGTGTGCGCGCGACGAGAAAGACTTCCGAGCGGTGCGAACCTCGAGGTAGACGTCAGCCTCCACACGCGCGCTGCTCTTTCTGAAGGCTCGCCAGCGTTGGAGCTCCATCTGCAGTCGGA

[illegible]

•

# FIGURE 18 - Page 2

ThrCysValThrGlnThrValAspPheSerLeuAspProThrPheThrIleGluThrIle  
 662 ACGTGTGTCACCCAGACAGTCGATTTTCAGCCTTGACCCTACCTTCACCATTGAGACAATC  
 TGCACACAGTGGGTCTGTCTAGCTAAAGTCGGAAGTGGGATGGAAGTGGTAACTCTGTAG  
  
 ThrLeuProGlnAspAlaValSerArgThrGlnArgArgGlyArgThrGlyArgGlyLys  
 722 ACGCTCCCCAAGATGCTGTCTCCCGCACTCAACGTGCGGGCAGGACTGGCAGGGGGAAG  
 TCGGAGGGGGTTCTACGACAGAGGGCGTGAGTTGCAGCCCCGTCTGACCGTCCCCCTTC  
  
 ProGlyIleTyrArgPheValAlaProGlyGluArgProSerGlyMetPheAspSerSer  
 782 CCAGGCATCTACAGATTTGTGGCACCGGGGAGCGCCCCCTCCGGCATGTTCTGACTCGTCC  
 GGTCCGTAGATGTCTAAACACCGTGGCCCCCTCGCGGGGAGGCCGTACAAGCTGAGCAGG  
 ^ ^  
 816 BGLI, 833 DRD1,  
  
 ValLeuCysGluCysTyrAspAlaGlyCysAlaTrpTyrGluLeuThrProAlaGluThr  
 842 GTCCTCTGTGAGTGCTATGACGCAGGCTGTGCTTGGTATGAGCTCACGCCCCGCGAGACT  
 CAGGAGACACTCACGATACTGCGTCCGACACGAACCATACTCGAGTGCGGGCGGCTCTGA  
 ^  
 881 SACI,  
  
 ThrValArgLeuArgAlaTyrMetAsnThrProGlyLeuProValCysGlnAspHisLeu  
 902 ACAGTTAGGCTACGAGCGTACATGAACACCCCGGGGCTTCCCGTGTGCCAGGACCATCTT  
 TGTCATCCGATGCTCGCATGTACTTGTGGGGCCCCGAAGGGCACACGGTCTCGGTAGAA  
 ^  
 931 SMAI XMAI,  
  
 GluPheTrpGluGlyValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGln  
 962 GAATTTTGGGAGGGCGTCTTTACAGGCCTCACTCATATAGATGCCCACTTCTATCCAG  
 CTTAAACCCCTCCCGCAGAAATGTCCGGAGTGAGTATATCTACGGGTGAAAGATAGGGTC  
 ^  
 985 STUI,  
  
 ThrLysGlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGlnAlaThrValCysAla  
 1022 ACAAGCAGAGTGGGGAGAACCTTCCTTACCTGGTAGCGTACCAAGCCACCGTGTGCGCT  
 TGTTTCGTCTCACCCCTCTTGAAGGAATGGACCATCGCATGGTTCGGTGGCACACGCGA  
 ^  
 1069 DRA3,  
  
 ArgAlaGlnAlaProProProSerTrpAspGlnMetTrpLysCysLeuIleArgLeuLys  
 1082 AGGGCTCAAGCCCCTCCCCATCGTGGGACCAGATGTGGAAGTGTGTTGATTCGCTCAAG  
 TCCCGAGTTCGGGGAGGGGGTAGCACCCCTGGTCTACACCTTCACAACTAAGCGGAGTTC  
  
 ProThrLeuHisGlyProThrProLeuLeuTyrArgLeuGlyAlaValGlnAsnGluIle  
 1142 CCCACCCTCCATGGGCCAACACCCCTGCTATACAGACTGGGCGCTGTTTCAAGTAAATC  
 GGGTGGGAGGTACCCGTTGTGGGGACGATATGTCTGACCCGCGACAAGTCTTACTTTAG  
 ^  
 1150 NCOI,  
  
 ThrLeuThrHisProValThrLysTyrIleMetThrCysMetSerAlaAspLeuGluVal  
 1202 ACCCTGACGCACCCAGTCACCAATACATCATGACATGCATGTCGGCCGACCTGGAGGTC  
 TGGGACTGCGTGGGTGAGTGGTTTATGTAGTACTGTACGTACAGCCGGCTGGACCTCCAG  
 ^ ^ ^ ^ ^  
 1230 BSPH1, 1234 DRD1, 1237 AVA3, 1245 EAG1 XMA3, 1250 DRD1,  
  
 ValThrSerThrTrpValLeuValGlyGlyValLeuAlaAlaLeuAlaAlaTyrCysLeu  
 1262 GTCACGAGCACCTGGGTGCTCGTTGGCGGCGTCTGGCTGCTTTGGCCGCGTATTGCCTG

SerThrGlyCysValValIleValGlyArgValValLeuSerGlyLysProAlaIleIle  
TCAACAGGCTGCGTGGTCATAGTGGGCAGGGTCGTCTTGTCCGGGAAGCCGGCAATCATA  
AGTTGTCCGACGCACCACTATCACCCGTCCCAGCAGAACAGGCCCTTCGGCCGTTAGTAT

ProAspArgGluValLeuTyrArgGluPheAspGluMetGluGluCysSerGlnHisLeu  
CCTGACAGGGAAGTCCTCTACCGAGAGTTTCATGAGATGGAAGAGTGCTCTCAGCACTTA  
GGACTGTCCCTTCAGGAGATGGCTCTCAAGCTACTCTACCTTCTCACGAGAGTCGTGAAT

ProTyrIleGluGlnGlyMetMetLeuAlaGluGlnPheLysGlnLysAlaLeuGlyLeu  
CCGTACATCGAGCAAGGGATGATGCTCGCCGAGCAGTTCAAGCAGAAGGCCCTCGGCCCTC  
GGCATGTAGCTCGTTCCCTACTACGAGCGGCTCGTCAAGTTCGTCTTCCGGGAGCCGGAG

LeuGlnThrAlaSerArgGlnAlaGluValIleAlaProAlaValGlnThrAsnTrpGln  
CTGCAGACCGCGTCCCGTCAGGCAGAGGTTATCGCCCCTGCTGTCCAGACCAACTGGCAA  
GACGCTCTGGCGCAGGGCAGTCCGTCTCCAATAGCGGGGACGACAGGTCTGGTTGACCGTT

LysLeuGluThrPheTrpAlaLysHisMetTrpAsnPheIleSerGlyIleGlnTyrLeu  
AAACTCGAGACCTTCTGGGCGAAGCATATGTGGAACCTTCATCAGTGGGATACAATACTTG  
TTTGAGCTCTGGAAGACCCGCTTCGTATACACCTTGAAGTAGTCACCCTATGTTATGAAC

AlaGlyLeuSerThrLeuProGlyAsnProAlaIleAlaSerLeuMetAlaPheThrAla  
GCGGGCTTGTC AACGCTGCCTGGTAACCCGCCATTGCTTCATTGATGGCTTTTACAGCT  
CGCCCGAACAGTTGCGACGGACCATTGGGGCGGTAACGAAGTAAC TACCGAAAATGTCGA

AlaValThrSerProLeuThrThrSerGlnThrLeuLeuPheAsnIleLeuGlyGlyTrp  
GCTGTCACCCAGCCCACTAACCACTAGCCAAACCCTCCTCTTCAACATATTGGGGGGGTGG  
CGACAGTGGTCGGGTGATTGGTGATCGGTTTGGGAGGAGAAGTTGTATAACCCCCCACC

ValAlaAlaGlnLeuAlaAlaProGlyAlaAlaThrAlaPheValGlyAlaGlyLeuAla  
GTGGCTGCCAGCTCGCCGCCCCGGTGCCGCTACTGCCTTTGTGGGCGCTGGCTTAGCT  
CACCGACGGGTCGAGCGGGCGGGGGCCACGGCGATGACGGAAACACCCGCGACCGAATCGA

GlyAlaAlaIleGlySerValGlyLeuGlyLysValLeuIleAspIleLeuAlaGlyTyr  
GGGCGCGCCATCGGCAGTGTTGGACTGGGGAAGGTCCTCATAGACATCCTTGCAGGGTAT  
CCGCGGCGGTAGCCGTCACAACCTGACCCCTTCCAGGAGTATCTGTAGGAACGTCCCAT

GlyAlaGlyValAlaGlyAlaLeuValAlaPheLysIleMetSerGlyGluValProSer  
GGCGCGGGCGTGGCGGGAGCTCTTGTGGCATTCAAGATCATGAGCGGTGAGGTCCCCTCC  
CCGCGCCCGCACCGCCCTCGAGAACACCGTAAGTTCTAGTACTCGCCACTCCAGGGGAGG

1878 SACI, 1899 BSPH1.

1922 ThrGluAspLeuValAsnLeuLeuProAlaIleLeuSerProGlyAlaLeuValValGly  
ACGGAGGACCTGGTCAATCTACTGCCCCCATCCTCTCGCCCGGAGCCCTCGTAGTCGGC  
TGCCTCCTGGACCAGTTAGATGACGGGCGGTAGGAGAGCGGGCCTCGGGAGCATCAGCCG  
1928 TTH3I,

1982 ValValCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaValGlnTrp  
GTGGTCTGTGCAGCAATACTGCGCCGGCACGTTGGCCCGGGCGAGGGGGCAGTGCAGTGG  
CACCAGACACGTCGTTATGACGCGGGCGTGCAACCGGGGCCGCTCCCCCGTCACGTCACC  
2004 NAEI, 2017 SMAI XMAI,

2042 MetAsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValSerProThrHisTyrVal  
ATGAACCGGCTGATAGCCTTCGCCTCCCGGGGAACCATGTTTCCCCACGCACTACGTG  
TACTTGGCCGACTATCGGAAGCGGAGGGCCCCCTTGGTACAAAGGGGGTGCCTGATGCAC  
2067 SMAI XMAI, 2093 DRA3,

2102 ProGluSerAspAlaAlaAlaArgValThrAlaIleLeuSerSerLeuThrValThrGln  
CCGGAGAGCGATGCAGCTGCCCCGCTCACTGCCATACTCAGCAGCCTCACTGTAACCCAG  
GGCCTCTCGCTACGTCGACGGGCGCAGTGACGGTATGAGTCGTCGGAGTGACATTGGGTC  
2115 PVU2, 2159 ALWN1,

2162 LeuLeuArgArgLeuHisGlnTrpIleSerSerGluCysThrThrProCysSerGlySer  
CTCCTGAGGCGACTGCACCACTGGATAAGCTCGGAGTGTAACCACTCCATGCTCCGGTTCC  
GAGGACTCCGCTGACGTGGTCACCTATTTCGAGCCTCACATGGTGAGGTACGAGGCCAAGG  
2164 MST2, 2220 ECON1,

2222 TrpLeuArgAspIleTrpAspTrpIleCysGluValLeuSerAspPheLysThrTrpLeu  
TGGCTAAGGGACATCTGGGACTGGATATGCGAGGTGTTGAGCGACTTTAAGACCTGGCTA  
ACCGATTCCCTGTAGACCCTGACCTATACGCTCCACAACCTCGCTGAAATTCTGGACCGAT

2282 LysAlaLysLeuMetProGlnLeuProGlyIleProPheValSerCysGlnArgGlyTyr  
AAAGCTAAGCTCATGCCACAGCTGCCTGGGATCCCCTTTGTGTCTGCCAGCGCGGGTAT  
TTTCGATTTCGAGTACGGTGTGACGGACCCTAGGGGAAACACAGGACGGTTCGCGCCCAT  
2285 ESP1, 2300 PVU2, 2310 BAMHI,

2342 LysGlyValTrpArgGlyAspGlyIleMetHisThrArgCysHisCysGlyAlaGluIle  
AAGGGGGTCTGGCGAGGGGACGGCATCATGCACACTCGCTGCCACTGTGGAGCTGAGATC  
TTCCCCCAGACCGCTCCCCTGCCGTAGTACGTGTGAGCGACGGTGACACCTCGACTCTAG

2402 ThrGlyHisValLysAsnGlyThrMetArgIleValGlyProArgThrCysArgAsnMet  
ACTGGACATGTCAAAAACGGGACGATGAGGATCGTCGGTCCTAGGACCTGCAGGAACATG  
TGACCTGTACAGTTTTTGCCTGCTACTCCTAGCAGCCAGGATCCTGGACGTCCTTGATAC  
2425 BSAB1, 2441 AVR2, 2448 SSE83871, 2449 PSTI,

2462 TrpSerGlyThrPheProIleAsnAlaTyrThrThrGlyProCysThrProLeuProAla  
TGGAGTGGGACCTTCCCCATTAATGCCTACACCACGGGGCCCTGTACCCCCCTTCTGCG  
ACCTCACCCCTGGAAGGGGTAATTACGGATGTGGTGCCCGGGACATGGGGGGAGGACGC  
2480 ASE1, 2497 APAI,

CGCTTCTCGCTACGTCGACGGGCGCAGTGACGGTATGAGTCGTCGGAGTGACATTGGGTC

# FIGURE 18 - Page 5

2522 ProAsnTyrThrPheAlaLeuTrpArgValSerAlaGluGluTyrValGluIleArgGln  
 CCGAACTACACGTTTCGCGCTATGGAGGGTGTCTGCAGAGGAATACGTGGAGATAAGGCAG  
 GGCTTGATGTGCAAGCGCGATACCTCCACAGACGTCTCCTTATGCACCTCTATTCCGTC  
 2553 PSTI,  
 2582 ValGlyAspPheHisTyrValThrGlyMetThrThrAspAsnLeuLysCysProCysGln  
 GTGGGGGACTTCCACTACGTGACGGGTATGACTACTGACAATCTTAAATGCCCGTGCCAG  
 CACCCCTGAAGGTGATGCACTGCCCATCTGATGACTGTTAGAATTTACGGGCACGGTC  
 2594 DRA3,  
 2642 ValProSerProGluPhePheThrGluLeuAspGlyValArgLeuHisArgPheAlaPro  
 GTCCCATCGCCGAATTTTTCACAGAATTGGACGGGGTGCGCCTACATAGGTTTGCGCC  
 CAGGGTAGCGGGCTTAAAAAGTGTCTTAACCTGCCCCACGCGGATGTATCCAAACGCGGG  
 2702 ProCysLysProLeuLeuArgGluGluValSerPheArgValGlyLeuHisGluTyrPro  
 CCCTGCAAGCCCTTGCTGCGGGAGGAGGTATCATTAGAGTAGGACTCCACGAATACCCG  
 GGGACGTTTCGGGAACGACGCCCTCCTCCATAGTAAGTCTCATCTGAGGTGCTTATGGGC  
 2757 HGIE2,  
 2762 ValGlySerGlnLeuProCysGluProGluProAspValAlaValLeuThrSerMetLeu  
 GTAGGGTCGCAATTACCTTGCGAGCCCGAACCGGACGTGGCCGTGTTGACGTCCATGCTC  
 CATCCAGCGTTAATGGAACGCTCGGGCTTGGCCTGCACCGGCACAACGAGGTACGAG  
 2809 AAT2,  
 2822 ThrAspProSerHisIleThrAlaGluAlaAlaGlyArgArgLeuAlaArgGlySerPro  
 ACTGATCCCTCCCATATAACAGCAGAGGCGGCCGGCGAAGGTTGGCGAGGGGATCACCC  
 TGACTAGGGAGGGTATATTGTCGTCTCCGCCGGCCGCTTCCAACCGCTCCCTAGTGGG  
 2850 EAG1 XMA3,  
 2882 ProSerValAlaSerSerSerAlaSerGlnLeuSerAlaProSerLeuLysAlaThrCys  
 CCCTCTGTGGCCAGCTCCTCGGCTAGCCAGCTATCCGCTCCATCTCTCAAGGCAACTTGC  
 GGGAGACACCGGTTCGAGGAGCCGATCGGTTCGATAGGCGAGGTAGAGAGTTCCGTTGAACG  
 2889 BALI, 2903 NHEI,  
 2942 ThrAlaAsnHisAspSerProAspAlaGluLeuIleGluAlaAsnLeuLeuTrpArgGln  
 ACCGCTAACCATGACTCCCTGATGCTGAGCTCATAGAGGCCAACCTCCTATGGAGGCAG  
 TGGCGATTGGTACTGAGGGGACTACGACTCGAGTATCTCCGTTGGAGGATACCTCCGTC  
 2966 ESP1, 2969 SACI,  
 3002 GluMetGlyGlyAsnIleThrArgValGluSerGluAsnLysValValIleLeuAspSer  
 GAGATGGGCGGCAACATCACCAGGGTTGAGTCAGAAAACAAAGTGGTGATTCTGGACTCC  
 CTCTACCCGCCGTTGTAGTGGTCCCAACTCAGTCTTTTGTTCACCACTAAGACCTGAGG  
 3062 PheAspProLeuValAlaGluGluAspGluArgGluIleSerValProAlaGluIleLeu  
 TTCGATCCGCTTGTGGCGGAGGAGGACGAGCGGGAGATCTCCGTACCCGCAGAAATCCTG  
 AAGCTAGGCGAACACCGCCTCCTCCTGCTCGCCCTCTAGAGGCATGGGCGTCTTTAGGAC  
 3096 BGL2,  
 ArgLysSerArgArgPheAlaGlnAlaLeuProValTrpAlaArgProAspTyrAsnPro

3122 CGGAAGTCTCGGAGATTTCGCCACGGCCCTGCCCGTTTGGGCGCGGCCGGAATAAACCCCGCCTTCAGAGCCTCTAAGCGGGTCCGGGACGGGCAAACCCGCGCCGCCTGATATTGGGG  
3143 ALWN1, 3164 EAG1 XMA3,  
ProLeuValGluThrTrpLysLysProAspTyrGluProProValValHisGlyCysPro  
3182 CCGCTAGTGGAGACGTGGAAAAAGCCCCGACTACGAACCACCTGTGGTCCATGGCTGCCCGGGCGATCACCTCTGCACCTTTTTTCGGGCTGATGCTTGGTGGACACCAGGTACCGACGGGC  
3217 HGIE2, 3229 NCOI,  
LeuProProProLysSerProProValProProProArgLysLysArgThrValValLeu  
3242 CTTCCACCTCCAAAGTCCCCCTCCTGTGCCTCCGCCTCGGAAGAAGCGGACGGTGGTCCTCGAAGGTGGAGTTTCAGGGGAGGACACGGAGGCGGAGCCTTCTTCGCCTGCCACCAGGAG  
3302 ThrGluSerThrLeuSerThrAlaLeuAlaGluLeuAlaThrArgSerPheGlySerSerACTGAATCAACCTATCTACTGCCTTGGCCGAGCTCGCCACCAGAAGCTTTGGCAGCTCTTGACTTAGTTGGGATAGATGACGGAACCGGCTCGAGCGGTGGTCTTCGAAACCGTCGAGG  
3332 SACI, 3346 HIND3,  
SerThrSerGlyIleThrGlyAspAsnThrThrThrSerSerGluProAlaProSerGly  
3362 TCAACTTCCGGCATTACGGGCGACAATACGACAACATCCTCTGAGCCCCGCCCTTCTGGCAGTTGAAGGCCGTAATGCCCCGCTGTTATGCTGTTGTAGGAGACTCGGGCGGGGAAGACCG  
3422 CysProProAspSerAspAlaGluSerTyrSerSerMetProProLeuGluGlyGluProTGCCCCCCCCGACTCCGACGCTGAGTCTATTCTCTCCATGCCCCCCTGGAGGGGGAGCCTACGGGGGGGCTGAGGCTGCGACTCAGGATAAGGAGGTACGGGGGGGACCTCCCCCTCGGA  
3437 EAM11051,  
GlyAspProAspLeuSerAspGlySerTrpSerThrValSerSerGluAlaAsnAlaGlu  
3482 GGGGATCCGGATCTTAGCGACGGGTCATGGTCAACGGTCAGTAGTGAGGCCAACGCGGAGCCCCTAGGCCTAGAATCGCTGCCCAGTACCAGTTGCCAGTCATCACTCCGGTTGCGCCTC  
3484 BAMHI, 3485 BSAB1, 3487 BSPE1,  
AspValValCysCysSerMetSerTyrSerTrpThrGlyAlaLeuValThrProCysAla  
3542 GATGTGCTGTGCTGCTCAATGTCTTACTCTTGACAGGCGCACTCGTCACCCCGTGCGCCCTACAGCACACGACGAGTTACAGAATGAGAACCTGTCCGCGTGAGCAGTGGGGCACGCGG  
3589 DRA3, 3600 SAC2,  
AlaGluGluGlnLysLeuProIleAsnAlaLeuSerAsnSerLeuLeuArgHisHisAsn  
3602 GCGGAAGAACAGAAACTGCCCATCAATGCACTAAGCAACTCGTTGCTACGTACCCACAATCGCCTTCTGTCTTTGACGGGTAGTTACGTGATTGTTGAGCAACGATGCAGTGGTGTTA  
3611 ALWN1, 3655 PFLM1,  
LeuValTyrSerThrThrSerArgSerAlaCysGlnArgGlnLysLysValThrPheAsp  
3662 TTGGTGTATTCCACCACCTCACGCAGTGCTTGCCAAAGGCAGAAAGATCACATTTGACAACACATAAGGTGGTGGAGTGCGTCACGAACGGTTTCCGTCTTCTTTCAGTGTAACCTG  
3681 DRA3,  
ArgLeuGlnValLeuAspSerHisTyrGlnAspValLeuLysGluValLysAlaAlaAla  
3722 AGACTGCAAGTTCTGGACAGCCATTACCAGGACGTACTCAAGGAGGTTAAAGCAGCGCGC



TCTGACGTTCAAGACCTGTCGGTAATGGTCCTGCATGAGTTCCTCCAATTTCTGTCGCGGC

3782 SerLysValLysAlaAsnLeuLeuSerValGluGluAlaCysSerLeuThrProProHis  
TCAAAGTGAAGGCTAACTTGCTATCCGTAGAGGAAGCTTGACGCTGACGCCCCACAC  
AGTTTTCACTCCGATTGAACGATAGGCATCTCCTTCGAACGTCGGACTGCGGGGGTGTG

3816 HIND3,

3842 SerAlaLysSerLysPheGlyTyrGlyAlaLysAspValArgCysHisAlaArgLysAla  
TCAGCCAAATCCAAGTTTGGTTATGGGGCAAAGACGTCGGTTGCCATGCCAGAAAGGCC  
AGTCGGTTTAGGTTCAAACCAATACCCCGTTTCTGCAGGCAACGGTACGGTCTTTCCGG

3875 AAT2, 3890 BGLI,

3902 ValThrHisIleAsnSerValTrpLysAspLeuLeuGluAspAsnValThrProIleAsp  
GTAACCCACATCAACTCCGTGTGGAAAGACCTTCTGGAAGACAATGTAACACCAATAGAC  
CATTGGGTGTAGTTGAGGCACACCTTCTGGAAGACCTTCTGTTACATTGTGGTTATCTG

3962 ThrThrIleMetAlaLysAsnGluValPheCysValGlnProGluLysGlyGlyArgLys  
ACTACCATCATGGCTAAGAACGAGGTTTTCTGCGTTTACGCTGAGAAGGGGGGTCGTAAG  
TGATGGTAGTACCGATTCTTGCTCCAAAAGACGCAAGTCGGACTCTTCCCCCAGCATTC

4022 ProAlaArgLeuIleValPheProAspLeuGlyValArgValCysGluLysMetAlaLeu  
CCAGCTCGTCTCATCGTGTCCCGATCTGGGCGTGCGCGTGTGCGAAAAGATGGCTTTG  
GGTCGAGCAGAGTAGCACAAAGGGGCTAGACCCGCACGCGCACACGCTTTTCTACCGAAAC

4082 TyrAspValValThrLysLeuProLeuAlaValMetGlySerSerTyrGlyPheGlnTyr  
TACGACGTGGTTACAAAGCTCCCTTGGCCGTGATGGGAAGCTCCTACGGATTCCAATAC  
ATGCTGCACCAATGTTTCGAGGGGAACCGGCACTACCCTTCGAGGATGCCTAAGGTTATG

4142 SerProGlyGlnArgValGluPheLeuValGlnAlaTrpLysSerLysLysThrProMet  
TCACCAGGACAGCGGGTTGAATTCCTCGTGCAAGCGTGGAAGTCCAAGAAAACCCCAATG  
AGTGGTCCTGTGCGCCCACTTAAGGAGCACGTTTCGCACCTTCAGGTTCTTTTGGGGTTAC

4160 ECORI,

4202 GlyPheSerTyrAspThrArgCysPheAspSerThrValThrGluSerAspIleArgThr  
GGGTTCTCGTATGATACCCGCTGCTTTGACTCCACAGTCACTGAGAGCGACATCCGTACG  
CCCAAGAGCATACTATGGGCGACGAACTGAGGTGTCAGTGACTCTCGCTGTAGGCATGC

4229 DRD1, 4236 ALWN1,

4262 GluGluAlaIleTyrGlnCysCysAspLeuAspProGlnAlaArgValAlaIleLysSer  
GAGGAGGCAATCTACCAATGTTGTGACCTCGACCCCCAAGCCCGCTGGCCATCAAGTCC  
CTCCTCCGTTAGATGGTTACAACACTGGAGCTGGGGGTTCCGGCGCACCGGTAGTTTACGG

4301 BGLI, 4308 BALI,

4322 LeuThrGluArgLeuTyrValGlyGlyProLeuThrAsnSerArgGlyGluAsnCysGly  
CTCACCGAGAGGCTTTATGTTGGGGGCCCTCTTACCAATTCAAGGGGGGAGAACTGCGGC  
GAGTGGCTCTCCGAAATACAACCCCGGAGAAATGGTTAAGTTCCCCCTCTTGACGCCG

4345 APAI,

4382 TyrArgArgCysArgAlaSerGlyValLeuThrThrSerCysGlyAsnThrLeuThrCys  
TATCGCAGGTGCCGCGGAGCGGCGTACTGACAACTAGCTGTGGTAACACCCTCACTTGC  
ATAGCGTCCACGGCGCGCTCGCCGCATGACTGTTGATCGACACCATTGTGGGAGTGAACG

002214 0472460

4442 TyrIleLysAlaArgAlaAlaCysArgAlaAlaGlyLeuGlnAspCysThrMetLeuVal  
 TACATCAAGGCCCGGGCAGCCTGTCTGAGCCGCGAGGGCTCCAGGACTGCACCATGCTCGTG  
 ATGTAGTTCCGGGGCCGTCGGACAGCTCGGCGTCCCGAGGTCCTGACGTGGTACGAGCAC  
 ^  
 4452 SMAI XMAI,  
 CysGlyAspAspLeuValValIleCysGluSerAlaGlyValGlnGluAspAlaAlaSer  
 4502 TGTGGCGACGACTTAGTCGTTATCTGTGAAAGCGCGGGGTCCAGGAGGACGCGGCGAGC  
 ACACCGCTGCTGAATCAGCAATAGACACTTTCGCGCCCCCAGGTCTCTCTGCGCCGCTCG  
 ^ ^  
 4508 DRD1, 4511 TTH3I,  
 LeuArgAlaPheThrGluAlaMetThrArgTyrSerAlaProProGlyAspProProGln  
 4562 CTGAGAGCCTTCACGGAGGCTATGACCAGGTAATCCGCCCCCTGGGGACCCCCACAA  
 GACTCTCGGAAGTGCCTCCGATACTGGTCCATGAGGCGGGGGGACCCCTGGGGGTGTT  
 ProGluTyrAspLeuGluLeuIleThrSerCysSerSerAsnValSerValAlaHisAsp  
 4622 CCAGAATACGACTTGGAGCTCATAACATCATGCTCCTCCAACGTGTCTGAGTCCGCCACGAC  
 GGTCTTATGCTGAACCTCGAGTATTGTAGTACGAGGAGGTTGCACAGTCAGCGGGTGTCT  
 ^  
 4637 SACI,  
 GlyAlaGlyLysArgValTyrTyrLeuThrArgAspProThrThrProLeuAlaArgAla  
 4682 GGCGCTGGAAAGAGGGTCTACTACCTCACCCGTGACCCTACAACCCCCCTCGCGAGAGCT  
 CCGCGACCTTCTCCCAGATGATGGAGTGGGCACTGGGATGTTGGGGGGAGCGCTCTCGA  
 ^  
 4731 NRUI,  
 AlaTrpGluThrAlaArgHisThrProValAsnSerTrpLeuGlyAsnIleIleMetPhe  
 4742 GCGTGGGAGACAGCAAGACACACTCCAGTCAATTCCTGGCTAGGCAACATAATCATGTTT  
 CGCACCTCTGTCTGTTCTGTGTGAGGTGAGTTAAGGACCGATCCGTTGTATTAGTACAAA  
 AlaProThrLeuTrpAlaArgMetIleLeuMetThrHisPhePheSerValLeuIleAla  
 4802 GCCCCACACTGTGGGCGAGGATGATACTGATGACCCATTCTTTAGCGTCTTATAGCC  
 CGGGGGTGTGACACCCGCTCCTACTATGACTACTGGGTAAAGAAATCGCAGGAATATCGG  
 ^ ^  
 4806 PFLM1, 4807 DRA3,  
 ArgAspGlnLeuGluGlnAlaLeuAspCysGluIleTyrGlyAlaCysTyrSerIleGlu  
 4862 AGGGACCAGCTTGAACAGGCCCTCGATTGCGAGATCTACGGGGCCTGCTACTCCATAGAA  
 TCCCTGGTCAACTTGTCCGGGAGCTAACGCTCTAGATGCCCCGGACGATGAGGTATCTT  
 ^  
 4893 BGL2,  
 ProLeuAspLeuProProIleIleGlnArgLeuHisGlyLeuSerAlaPheSerLeuHis  
 4922 CCACTGGATCTACCTCCAATCATTCAAAGACTCCATGGCCTCAGCGCATTTTCACTCCAC  
 GGTGACCTAGATGGAGGTTAGTAAGTTTCTGAGGTACCGGAGTCGCGTAAAAGTGAGGTG  
 ^  
 4954 NCOI,  
 SerTyrSerProGlyGluIleAsnArgValAlaAlaCysLeuArgLysLeuGlyValPro  
 4982 AGTTACTCTCCAGGTGAAATCAATAGGGTGGCCGCATGCCTCAGAAACTTGGGGTACCG  
 TCAATGAGAGGTCCACTTTAGTTATCCACCGGCGTACGGAGTCTTTTGAACCCCATGGC  
 ^ ^  
 5015 SPHI, 5035 KPNI,

002440 047260

# FIGURE 18 - Page 9

5042 ProLeuArgAlaTrpArgHisArgAlaArgSerValArgAlaArgLeuLeuAlaArgGly  
CCCTTGGCAGCTTGGAGACACCGGGCCCGGAGCGTCCGCGCTAGGCTTCTGGCCAGAGGA  
GGGAACGCTCGAACCTCTGTGGCCCGGGCTCGCAGGCGCGATCCGAAGACCGGTCTCCT

5064 APAI, 5091 BALI,

5102 GlyArgAlaAlaIleCysGlyLysTyrLeuPheAsnTrpAlaValArgThrLysLeuLys  
GGCAGGGCTGCCATATGTGGCAAGTACCTCTTCAACTGGGCAGTAAGAACAAAGCTCAAA  
CCGTCCCGACGGTATACACCGTTCATGGAGAAGTTGACCCGTCATTCTTGTTCGAGTTT

5113 NDEI,

5162 LeuThrProIleAlaAlaAlaGlyGlnLeuAspLeuSerGlyTrpPheThrAlaGlyTyr  
CTCACTCCAATAGCGGCGCTGGCCAGCTGGACTTGTCGGCTGGTTCACGGCTGGCTAC  
GAGTGAGGTTATCGCCGGCGACCGGTGACCTGAACAGGCCGACCAAGTGCCGACCGATG

5174 NOTI, 5175 EAG1 XMA3, 5182 BALI, 5186 PVU2,

5222 SerGlyGlyAspIleTyrHisSerValSerHisAlaArgProArgTrpIleTrpPheCys  
AGCGGGGAGACATTTATCACAGCGTGTCTCATGCCCGGGCCCGCTGGATCTGGTTTTGC  
TCGCCCCCTCTGTAAATAGTGTGCGACAGAGTACGGGCGGGGGCGACCTAGACCAAACG

5240 DRA3,

5282 LeuLeuLeuLeuAlaAlaGlyValGlyIleTyrLeuLeuProAsnArgMetSerThrAsn  
CTACTCCTGCTTGCTGCAGGGGTAGGCATCTACCTCCTCCCAACCGAATGAGCACGAAT  
GATGAGGACGAACGACGTCCCCATCCGTAGATGGAGGAGGGTTGGCTTACTCGTGCTTA

5295 PSTI,

5342 ProLysProGlnArgLysThrLysArgAsnThrAsnArgArgProGlnAspValLysPhe  
CCTAAACCTCAAAGAAAGACCAACGTAACACCAACCGGCGGCCGAGGACGTCAAGTTC  
GGATTGGAGTTTCTTTCTGGTTTGCAATTGTGGTTGGCCCGCGCGTCTGCAGTTCAAG

5380 NOTI, 5381 EAG1 XMA3, 5390 AAT2, 5401 SMAI XMAI,

5402 ProGlyGlyGlyGlnIleValGlyGlyValTyrLeuLeuProArgArgGlyProArgLeu  
CCGGGTGGCGGTGAGATCGTTGGTGGAGTTTACTTGTTGCCGCGCAGGGGCCCTAGATTG  
GGCCACCGCCAGTCTAGCAACCACCTCAAATGAACAACGGCGCGTCCCCGGGATCTAAC

5449 APAI,

5462 GlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGlyArgArgGlnPro  
GGTGTGCGCGCGACGAGAAAGACTTCCGAGCGGTGCAACCTCGAGGTAGACGTACGCCT  
CCACACGCGCGCTGCTCTTTCTGAAGGCTCGCCAGCGTTGGAGCTCCATCTGCAGTCGGA

5467 BSSH2, 5478 XMNI, 5502 XHOI, 5511 AAT2,

5522 IleProLysAlaArgArgProGluGlyArgThrTrpAlaGlnProGlyTyrProTrpPro  
ATCCCCAAGGCTCGTCGGCCCGAGGGCAGGACCTGGGCTCAGCCCGGGTACCCTTGGCCC  
TAGGGGTTCCGAGCAGCCGGGCTCCCGTCTGGACCCGAGTCGGGCCCATGGGAACCGGG

5548 ALWN1, 5558 ESP1, 5564 SMAI XMAI, 5568 KPNI,

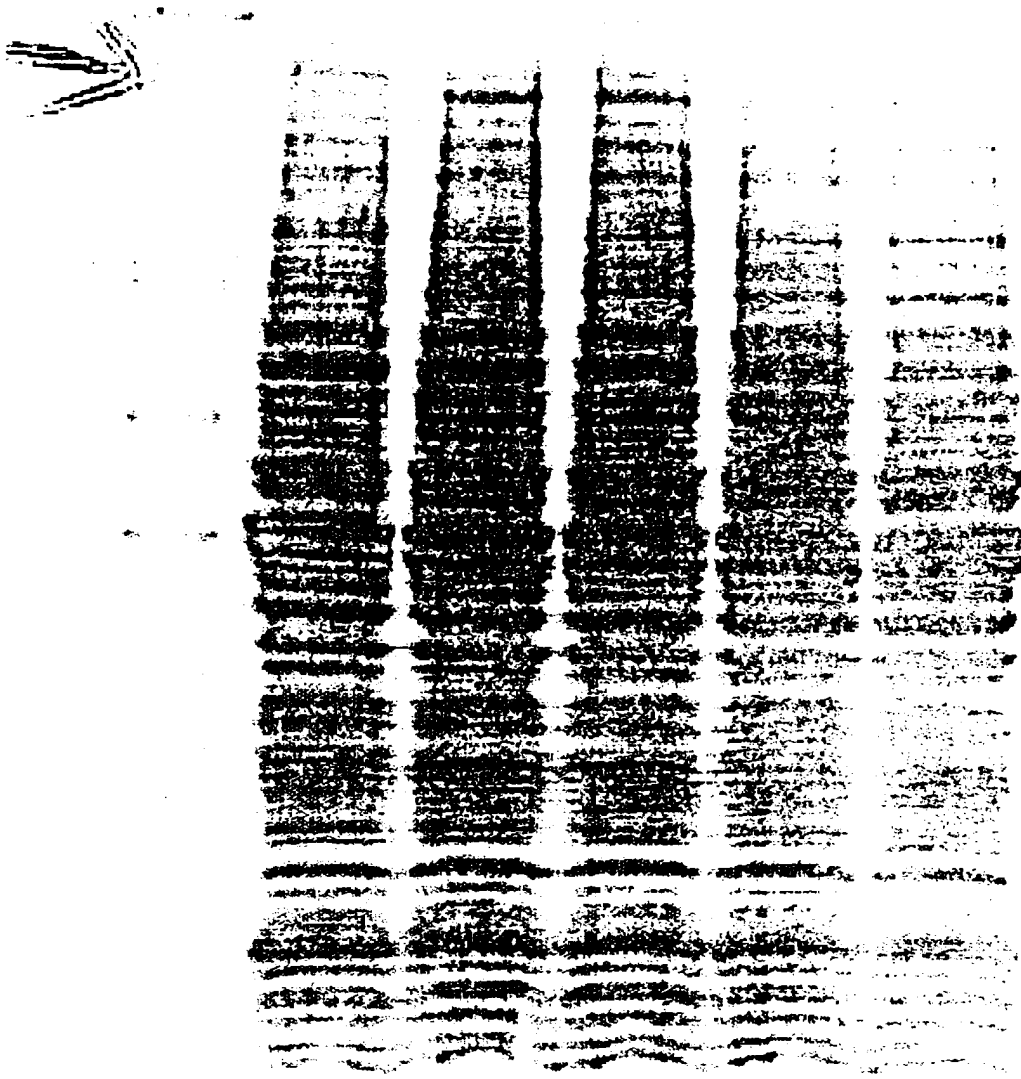
5582 LeuTyrGlyAsnGluGlyCysGlyTrpAlaGlyTrpLeuLeuSerProArgGlySerArg  
CTCTATGGCAATGAGGGCTGCGGGTGGGCGGGATGGCTCCTGTCTCCCCGTGGCTCTCGG  
GAGATACCGTTACTCCCGACGCCCCACCCGCCCTACCGAGGACAGAGGGGCACCGAGAGCC

5650 APAI, 5696 CLAI,

5724 HGIE2, 5750 KAS1 NARI, 5756 ECON1,

5772 BSTXI, 5775 APAI,

5854 SALI,

[illegible]

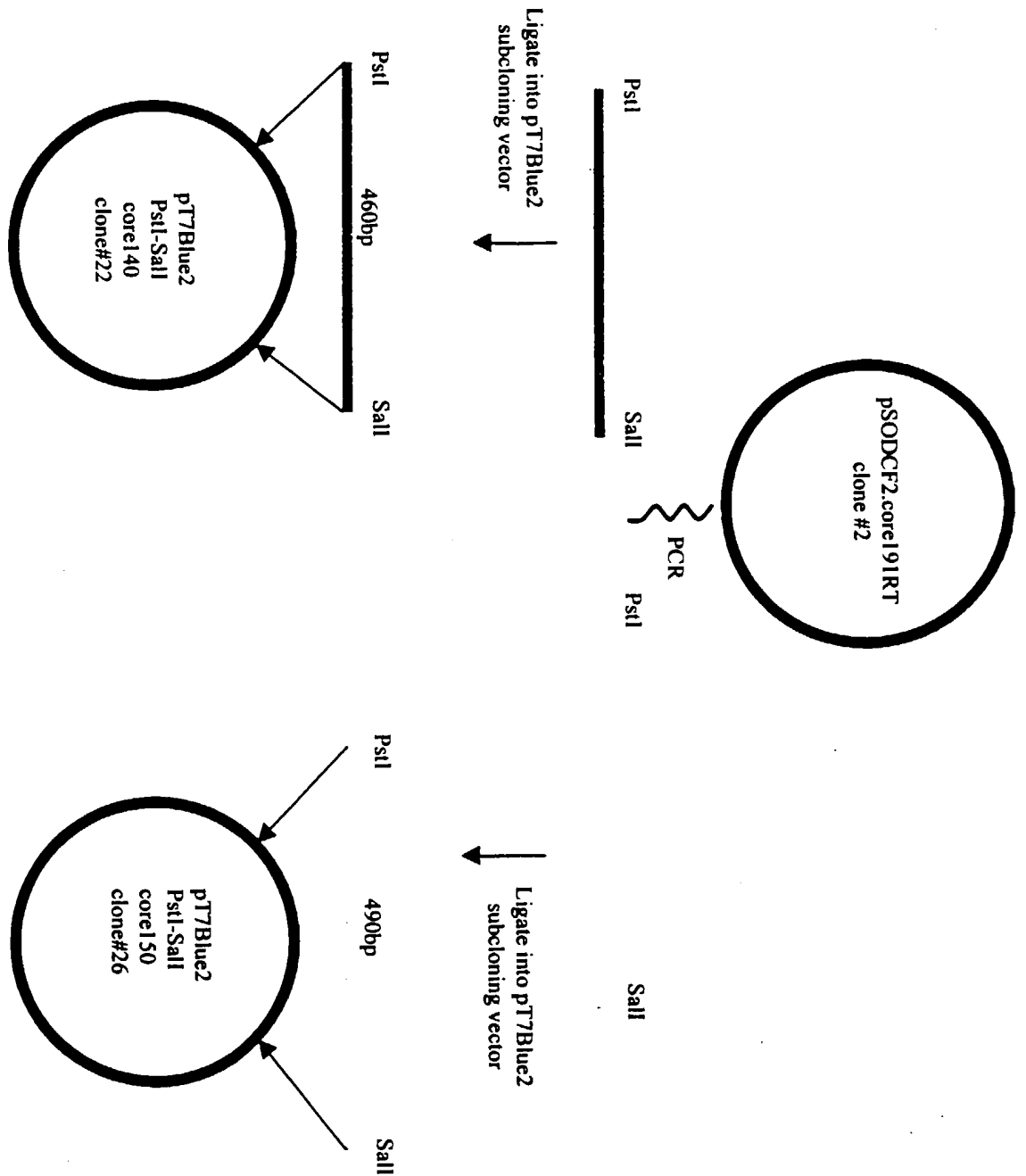
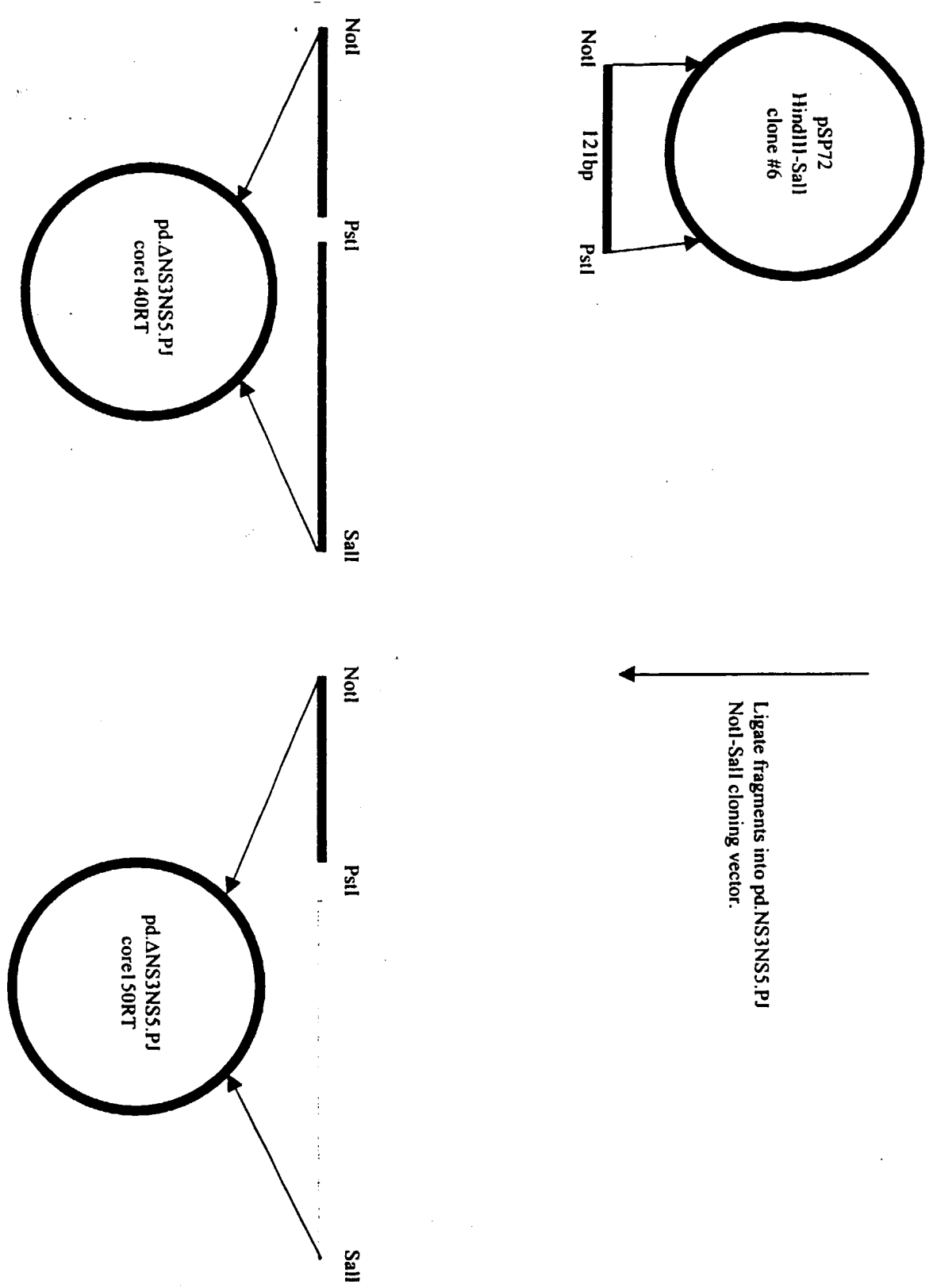


FIGURE 20 - Pa



# FIGURE 21 - Page 1

MetAlaAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsn  
2 AGCTTACAAAACAAAATGGCTGCATATGCAGCTCAGGGCTATAAGGTGCTAGTACTCAAC  
TCGAATGTTTTGTTTTACCGACGTATACGTTCGAGTCCCGATATTCCACGATCATGAGTTG  
^ ^ ^  
1 HIND3, 24 NDEI, 52 SCAI,  
ProSerValAlaAlaThrLeuGlyPheGlyAlaTyrMetSerLysAlaHisGlyIleAsp  
62 CCCTCTGTTGCTGCAACACTGGGCTTTGGTGCTTACATGTCCAAGGCTCATGGGATCGAT  
GGGAGACAACGACGTTGTGACCCGAAACCACGAATGTACAGGTTCCGAGTACCCCTAGCTA  
^  
116 CLAI,  
ProAsnIleArgThrGlyValArgThrIleThrThrGlySerProIleThrTyrSerThr  
122 CCTAACATCAGGACCGGGGTGAGAACAAATTACCACTGGCAGCCCCATCAGTACTCCACC  
GGATTGTAGTCCTGGCCCCACTCTTGTTAATGGTGACCGTCGGGGTAGTGCATGAGGTGG  
TyrGlyLysPheLeuAlaAspGlyGlyCysSerGlyGlyAlaTyrAspIleIleIleCys  
182 TACGGCAAGTTCCTTGCCGACGGCGGGTGCTCGGGGGGCGCTTATGACATAATAATTTGT  
ATGCCGTTCAAGGAACGGCTGCCGCCCCACGAGCCCCCGGAATACTGTATTATTAAACA  
AspGluCysHisSerThrAspAlaThrSerIleLeuGlyIleGlyThrValLeuAspGln  
242 GACGAGTGCCACTCCACGGATGCCACATCCATCTTGGGCATTGGCACTGTCCTTGACCAA  
CTGCTCACGGTGAGGTGCCTACGGTGTAGGTAGAACCCGTAACCGTGACAGGAACTGGTT  
AlaGluThrAlaGlyAlaArgLeuValValLeuAlaThrAlaThrProProGlySerVal  
302 GCAGAGACTGCGGGGGCGAGACTGGTTGTGCTCGCCACCGCCACCCCTCCGGGCTCCGTC  
CGTCTCTGACGCCCCCGCTCTGACCAACACGAGCGGTGGCGGTGGGGAGGCCCGAGGCAG  
^  
303 ALWN1,  
ThrValProHisProAsnIleGluGluValAlaLeuSerThrThrGlyGluIleProPhe  
362 ACTGTGCCCCATCCCAACATCGAGGAGGTTGCTCTGTCCACCACCGAGAGATCCCTTTT  
TGACACGGGGTAGGGTTGTAGCTCCTCCAACGAGACAGGTGGTGGCCTCTCTAGGGAAA  
TyrGlyLysAlaIleProLeuGluValIleLysGlyGlyArgHisLeuIlePheCysHis  
422 TACGGCAAGGCTATCCCCCTCGAAGTAATCAAGGGGGGGAGACATCTCATCTTCTGTTCAT  
ATGCCGTTCCGATAGGGGGAGCTTCATTAGTTCCCCCCTCTGTAGAGTAGAAGACAGTA

00347-647260



| 2017-2018 |  | 2018-2019 |  | 2019-2020 |  | 2020-2021 |  | 2021-2022 |  | 2022-2023 |  | 2023-2024 |  | 2024-2025 |  | 2025-2026 |  | 2026-2027 |  | 2027-2028 |  | 2028-2029 |  | 2029-2030 |  | 2030-2031 |  | 2031-2032 |  | 2032-2033 |  | 2033-2034 |  | 2034-2035 |  | 2035-2036 |  | 2036-2037 |  | 2037-2038 |  | 2038-2039 |  | 2039-2040 |  | 2040-2041 |  | 2041-2042 |  | 2042-2043 |  | 2043-2044 |  | 2044-2045 |  | 2045-2046 |  | 2046-2047 |  | 2047-2048 |  | 2048-2049 |  | 2049-2050 |  | 2050-2051 |  | 2051-2052 |  | 2052-2053 |  | 2053-2054 |  | 2054-2055 |  | 2055-2056 |  | 2056-2057 |  | 2057-2058 |  | 2058-2059 |  | 2059-2060 |  | 2060-2061 |  | 2061-2062 |  | 2062-2063 |  | 2063-2064 |  | 2064-2065 |  | 2065-2066 |  | 2066-2067 |  | 2067-2068 |  | 2068-2069 |  | 2069-2070 |  | 2070-2071 |  | 2071-2072 |  | 2072-2073 |  | 2073-2074 |  | 2074-2075 |  | 2075-2076 |  | 2076-2077 |  | 2077-2078 |  | 2078-2079 |  | 2079-2080 |  | 2080-2081 |  | 2081-2082 |  | 2082-2083 |  | 2083-2084 |  | 2084-2085 |  | 2085-2086 |  | 2086-2087 |  | 2087-2088 |  | 2088-2089 |  | 2089-2090 |  | 2090-2091 |  | 2091-2092 |  | 2092-2093 |  | 2093-2094 |  | 2094-2095 |  | 2095-2096 |  | 2096-2097 |  | 2097-2098 |  | 2098-2099 |  | 2099-2100 |  | 2100-2101 |  | 2101-2102 |  | 2102-2103 |  | 2103-2104 |  | 2104-2105 |  | 2105-2106 |  | 2106-2107 |  | 2107-2108 |  | 2108-2109 |  | 2109-2110 |  | 2110-2111 |  | 2111-2112 |  | 2112-2113 |  | 2113-2114 |  | 2114-2115 |  | 2115-2116 |  | 2116-2117 |  | 2117-2118 |  | 2118-2119 |  | 2119-2120 |  | 2120-2121 |  | 2121-2122 |  | 2122-2123 |  | 2123-2124 |  | 2124-2125 |  | 2125-2126 |  | 2126-2127 |  | 2127-2128 |  | 2128-2129 |  | 2129-2130 |  | 2130-2131 |  | 2131-2132 |  | 2132-2133 |  | 2133-2134 |  | 2134-2135 |  | 2135-2136 |  | 2136-2137 |  | 2137-2138 |  | 2138-2139 |  | 2139-2140 |  | 2140-2141 |  | 2141-2142 |  | 2142-2143 |  | 2143-2144 |  | 2144-2145 |  | 2145-2146 |  | 2146-2147 |  | 2147-2148 |  | 2148-2149 |  | 2149-2150 |  | 2150-2151 |  | 2151-2152 |  | 2152-2153 |  | 2153-2154 |  | 2154-2155 |  | 2155-2156 |  | 2156-2157 |  | 2157-2158 |  | 2158-2159 |  | 2159-2160 |  | 2160-2161 |  | 2161-2162 |  | 2162-2163 |  | 2163-2164 |  | 2164-2165 |  | 2165-2166 |  | 2166-2167 |  | 2167-2168 |  | 2168-2169 |  | 2169-2170 |  | 2170-2171 |  | 2171-2172 |  | 2172-2173 |  | 2173-2174 |  | 2174-2175 |  | 2175-2176 |  | 2176-2177 |  | 2177-2178 |  | 2178-2179 |  | 2179-2180 |  | 2180-2181 |  | 2181-2182 |  | 2182-2183 |  | 2183-2184 |  | 2184-2185 |  | 2185-2186 |  | 2186-2187 |  | 2187-2188 |  | 2188-2189 |  | 2189-2190 |  | 2190-2191 |  | 2191-2192 |  | 2192-2193 |  | 2193-2194 |  | 2194-2195 |  | 2195-2196 |  | 2196-2197 |  | 2197-2198 |  | 2198-2199 |  | 2199-2200 |  | 2200-2201 |  | 2201-2202 |  | 2202-2203 |  | 2203-2204 |  | 2204-2205 |  | 2205-2206 |  | 2206-2207 |  | 2207-2208 |  | 2208-2209 |  | 2209-2210 |  | 2210-2211 |  | 2211-2212 |  | 2212-2213 |  | 2213-2214 |  | 2214-2215 |  | 2215-2216 |  | 2216-2217 |  | 2217-2218 |  | 2218-2219 |  | 2219-2220 |  | 2220-2221 |  | 2221-2222 |  | 2222-2223 |  | 2223-2224 |  | 2224-2225 |  | 2225-2226 |  | 2226-2227 |  | 2227-2228 |  | 2228-2229 |  | 2229-2230 |  | 2230-2231 |  | 2231-2232 |  | 2232-2233 |  | 2233-2234 |  | 2234-2235 |  | 2235-2236 |  | 2236-2237 |  | 2237-2238 |  | 2238-2239 |  | 2239-2240 |  | 2240-2241 |  | 2241-2242 |  | 2242-2243 |  | 2243-2244 |  |
|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|
|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|-----------|--|

TCCCGAGTTCGGGGAGGGGGTAGCACCTGGTCTACACCTTCACAACTAAGCGGAGTTC

ProThrLeuHisGlyProThrProLeuLeuTyrArgLeuGlyAlaValGlnAsnGluIle  
 1142 CCCACCCTCCATGGGCCAACACCCCTGCTATACAGACTGGGCGCTGTTTCAAGATGAAATC  
 GGGTGGGAGGTACCCGGTTGTGGGGACGATATGTCTGACCCGCGACAAGTCTTACTTTAG  
 ^  
 1150 NCOI,

ThrLeuThrHisProValThrLysTyrIleMetThrCysMetSerAlaAspLeuGluVal  
 1202 ACCCTGACGCACCCAGTCACCAAATACATCATGACATGCATGTCGGCCGACCTGGAGGTC  
 TGGGACTGCGTGGGTGAGTGGTTTATGTAGTACTGTACGTACAGCCGGCTGGACCTCCAG  
 ^ ^ ^ ^ ^  
 1230 BSPH1, 1234 DRD1, 1237 AVA3, 1245 EAG1 XMA3, 1250 DRD1,

ValThrSerThrTrpValLeuValGlyGlyValLeuAlaAlaLeuAlaAlaTyrCysLeu  
 1262 GTCACGAGCACCTGGGTGCTCGTTGGCGGCGTCCTGGCTGCTTTGGCCGCGTATTGCCTG  
 CAGTGTCTGCTGGACCCACGAGCAACCGCCGAGGACCGACGAAACCGGCGCATAACGGAC

SerThrGlyCysValValIleValGlyArgValValLeuSerGlyLysProAlaIleIle  
 1322 TCAACAGGCTGCGTGGTTCATAGTGGGCAGGGTCGTCTTGTCCGGGAAGCCGGCAATCATA  
 AGTTGTCCGACGCACAGTATCACCCGTCCAGCAGAACAGGCCCTTCGGCCGTTAGTAT  
 ^  
 1369 NAEI,

ProAspArgGluValLeuTyrArgGluPheAspGluMetGluGluCysSerGlnHisLeu  
 1382 CCTGACAGGGAAGTCTCTACCGAGAGTTCGATGAGATGGAAGAGTGCTCTCAGCACTTA  
 GGACTGTCCCTTCAGGAGATGGCTCTCAAGCTACTCTACCTTCTCAGGAGAGTCGTGAAT  
 ^  
 1385 DRD1,

ProTyrIleGluGlnGlyMetMetLeuAlaGluGlnPheLysGlnLysAlaLeuGlyLeu  
 1442 CCGTACATCGAGCAAGGGATGATGCTCGCCGAGCAGTTCAGCAGAAAGGCCCTCGGCCTC  
 GGCATGTAGCTCGTTCCCTACTACGAGCGGCTCGTCAAGTTCGTCTTCCGGGAGCCGGAG

LeuGlnThrAlaSerArgGlnAlaGluValIleAlaProAlaValGlnThrAsnTrpGln  
 1502 CTGCAGACCGCGTCCCGTCAGGCAGAGGTTATCGCCCCTGCTGTCCAGACCAACTGGCAA  
 GACGTCTGGCGCAGGGCAGTCCGTCTCCAATAGCGGGGACGACAGGTCTGGTTGACCGTT  
 ^ ^  
 1502 PSTI, 1507 TTH3I,

LysLeuGluThrPheTrpAlaLysHisMetTrpAsnPheIleSerGlyIleGlnTyrLeu  
 1562 AAATCGAGACCTTCTGGGCGAAGCATATGTGGAACCTTCATCAGTGGGATACAATACTTG  
 TTTGAGCTCTGGAAGACCCGTTCTGATACACCTTGAAGTAGTCACCCTATGTTATGAAC  
 ^ ^  
 1565 XHOI, 1586 NDEI,

AlaGlyLeuSerThrLeuProGlyAsnProAlaIleAlaSerLeuMetAlaPheThrAla  
 1622 GCGGGCTTGTCAACGCTGCCTGGTAACCCCGCCATTGCTTCATTGATGGCTTTTACAGCT  
 CGCCCGAACAGTTGCGACGGACCATTGGGGCGGTAACGAAGTAACTACCGAAAATGTCTGA  
 ^ ^  
 1643 BSTE2, 1677 ALWN1 PVU2,

AlaValThrSerProLeuThrThrSerGlnThrLeuLeuPheAsnIleLeuGlyGlyTrp  
 1682 GCTGTACACGACCCACTAACCCTAGCCAAACCCCTCCTCTTCAACATATTGGGGGGGTGG  
 CGACAGTGGTGGGTGATTGGTGATCGGTTTGGGAGGAGAAGTTGTATAACCCCCCACC

097249-1330

ValAlaAlaGlnLeuAlaAlaProGlyAlaAlaThrAlaPheValGlyAlaGlyLeuAla  
1742 GTGGCTGCCCAGCTCGCCGCCCCCGGTGCCGCTACTGCCTTTGTGGGCGCTGGCTTAGCT  
CACCGACGGGTCGAGCGGCGGGGGCCACGGCGATGACGGAAACACCCGCGACCGAATCGA  
1794 ESP1,  
GlyAlaAlaIleGlySerValGlyLeuGlyLysValLeuIleAspIleLeuAlaGlyTyr  
1802 GGCGCCGCCATCGGCAGTGTGGACTGGGGAAGGTCCTCATAGACATCCTTGCAGGGTAT  
CCGCGGCGGTAGCCGTCACAACCTGACCCCTTCCAGGAGTATCTGTAGGAACGTCCCAT  
1802 KAS1 NARI,  
GlyAlaGlyValAlaGlyAlaLeuValAlaPheLysIleMetSerGlyGluValProSer  
1862 GGCGCGGGCGTGGCGGGAGCTCTTGTGGCATTCAAGATCATGAGCGGTGAGGTCCCCCTCC  
CCGCGCCCCGACCGCCCTCGAGAACACCGTAAGTTCTAGTACTCGCCACTCCAGGGGAGG  
1878 SACI, 1899 BSPH1,  
ThrGluAspLeuValAsnLeuLeuProAlaIleLeuSerProGlyAlaLeuValValGly  
1922 ACGGAGGACCTGGTCAATCTACTGCCCGCCATCCTCTCGCCCGGAGCCCTCGTAGTCGGC  
TGCCTCTCTGGACCAGTTAGATGACGGGCGGTAGGAGAGCGGGCCTCGGGAGCATCAGCCG  
1928 TTH3I,  
ValValCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaValGlnTrp  
1982 GTGGTCTGTGCAGCAATACTGCGCCGGCACGTTGGCCCGGGCGAGGGGGCAGTGCAGTGG  
CACCAGACACGTCGTTATGACGCGGGCGGTGCAACCGGGCCCGCTCCCCCGTCACGTCACC  
2004 NAEI, 2017 SMAI XMAI,  
MetAsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValSerProThrHisTyrVal  
2042 ATGAACCGGCTGATAGCCTTCGCCTCCCGGGGAACCATGTTTCCCCCAGCCTACTAGTG  
TACTTGGCCGACTATCGGAAGCGGAGGGCCCCCTTGGTACAAAGGGGTGCGTGATGCAC  
2067 SMAI XMAI, 2093 DRA3,  
ProGluSerAspAlaAlaAlaArgValThrAlaIleLeuSerSerLeuThrValThrGln  
2102 CCGGAGAGCGATGCAGCTGCCCGCGTCACTGCCATACTCAGCAGCCTCACTGTAACCCAG  
GGCCTCTCGCTACGTCGACGGGCGCAGTGACGGTATGAGTCGTCGGAGTGACATTGGGT  
2115 PVU2, 2159 ALWN1,  
LeuLeuArgArgLeuHisGlnTrpIleSerSerGluCysThrThrProCysSerGlySer  
2162 CTCCTGAGGCGACTGCACCACTGGATAAGCTCGGAGTGTAACACTCCATGCTCCGGTTCC  
GAGGACTCCGCTGACGTGGTACCTATTTCGAGCCTCACATGGTGAGGTACGAGGCCAAGG  
2164 MST2, 2220 ECON1,  
TrpLeuArgAspIleTrpAspTrpIleCysGluValLeuSerAspPheLysThrTrpLeu  
2222 TGGCTAAGGGACATCTGGGACTGGATATGCGAGGTGTGAGCGACTTTAAGACCTGGCTA  
ACCGATTCCCTGTAGACCCTGACCTATACGCTCCACAACTCGCTGAAATTCTGGACCGAT  
2282 LysAlaLysLeuMetProGlnLeuProGlyIleProPheValSerCysGlnArgGlyTyr  
AAAGCTAAGCTCATGCCACAGCTGCCTGGGATCCCCTTTGTGTCCTGCCAGCGCGGGTAT  
TTTCGATTTCGAGTACGGTGTGACGGACCCTAGGGGAAACACAGGACGGTCGCGCCCAT  
2285 ESP1, 2300 PVU2, 2310 BAMHI,

LysGlyValTrpArgGlyAspGlyIleMetHisThrArgCysHisCysGlyAlaGluIle  
 2342 AAGGGGGTCTGGCGAGGGGACGGCATCATGCACACTCGCTGCCACTGTGGAGCTGAGATC  
 TTCCCCAGACCGCTCCCCTGCCGTAGTACGTGTGAGCGACGGTGACACCTCGACTCTAG  
  
 ThrGlyHisValLysAsnGlyThrMetArgIleValGlyProArgThrCysArgAsnMet  
 2402 ACTGGACATGTCAAAAACGGGACGATGAGGATCGTCGGTCCTAGGACCTGCAGGAACATG  
 TGACCTGTACAGTTTTTGCCCTGCTACTCCTAGCAGCCAGGATCCTGGACGTCCTTGATC  
 ^ ^ ^  
 2425 BSAB1, 2441 AVR2, 2448 SSE83871, 2449 PSTI,  
  
 TrpSerGlyThrPheProIleAsnAlaTyrThrThrGlyProCysThrProLeuProAla  
 2462 TGGAGTGGGACCTTCCCCATTAATGCCTACACCACGGGCCCCTGTACCCCCCTTCCTGCG  
 ACCTCACCTGGAAGGGGTAAATTACGGATGTGGTGCCCGGGGACATGGGGGGAAGGACGC  
 ^ ^  
 2480 ASE1, 2497 APAI,  
  
 ProAsnTyrThrPheAlaLeuTrpArgValSerAlaGluGluTyrValGluIleArgGln  
 2522 CCGAACTACACGTTTCGCGCTATGGAGGGTGTCTGCAGAGGAATACGTGGAGATAAGGCAG  
 GGCTTGATGTGCAAGCGCGATACCTCCCACAGACGTCTCCTTATGCACCTCTATTCCGTC  
 ^  
 2553 PSTI,  
  
 ValGlyAspPheHisTyrValThrGlyMetThrThrAspAsnLeuLysCysProCysGln  
 2582 GTGGGGGACTTCCACTACGTGACGGGTATGACTACTGACAATCTTAAATGCCCCGTGCCAG  
 CCCCCCTGAAGGTGATGCACTGCCCATAGTACTGATGACTGTTAGAATTTACGGGCACGGTC  
 ^  
 2594 DRA3,  
  
 ValProSerProGluPhePheThrGluLeuAspGlyValArgLeuHisArgPheAlaPro  
 2642 GTCCCATCGCCCGAATTTTTACAGAATTGGACGGGTGCGCCTACATAGGTTTGCGCCC  
 CAGGGTAGCGGGCTTAAAAAGTGTCTTAACCTGCCCCACGCGGATGTATCCAAACGCGGG  
  
 ProCysLysProLeuLeuArgGluGluValSerPheArgValGlyLeuHisGluTyrPro  
 2702 CCCTGCAAGCCCTTGCTGCGGGAGGAGGTATCATTGAGAGTAGGACTCCACGAATACCCG  
 GGGACGTTTCGGGAACGACGCCCTCCTCCATAGTAAGTCTCATCCTGAGGTGCTTATGGGC  
 ^  
 2757 HGIE2,  
  
 ValGlySerGlnLeuProCysGluProGluProAspValAlaValLeuThrSerMetLeu  
 2762 GTAGGGTCGCAATTACCTTGCGAGCCCGAACCGGACGTGGCCGTGTTGACGTCCATGCTC  
 CATCCCAGCGTTAATGGAACGCTCGGGCTTGGCCTGCACCGGCACAAGTGCAGGTACGAG  
 ^  
 2809 AAT2,  
  
 ThrAspProSerHisIleThrAlaGluAlaAlaGlyArgArgLeuAlaArgGlySerPro  
 2822 ACTGATCCCTCCCATATAACAGCAGAGGCGCGCGGCGAAGGTTGGCGAGGGGATCACCC  
 TGACTAGGGAGGGTATATTGTCGTCTCCGCGGCGCCGCTTCCAACCGCTCCCTAGTGGG  
 ^  
 2850 EAG1 XMA3,  
  
 ProSerValAlaSerSerSerAlaSerGlnLeuSerAlaProSerLeuLysAlaThrCys  
 2882 CCCTCTGTGGCCAGCTCCTCGGCTAGCCAGCTATCCGCTCCATCTCTCAAGGCAACTTGC  
 GGGAGACACCGGTGAGGAGCCGATCGGTGATAGGCGAGGTAGAGAGTTCCGTTGAACG  
 ^  
 2889 BALI, 2903 NHEI,

2342  
 2402  
 2462  
 2522  
 2582  
 2642  
 2702  
 2762  
 2822  
 2882

2942 ThrAlaAsnHisAspSerProAspAlaGluLeuIleGluAlaAsnLeuLeuTrpArgGln  
ACCGCTAACCATGACTCCCCTGATGCTGAGCTCATAGAGGCCAACCTCCTATGGAGGCAG  
TGGCGATTGGTACTGAGGGGACTACGACTCGAGTATCTCCGGTTGGAGGATACCTCCGTC  
^ ^

2966 ESP1, 2969 SACI,

3002 GluMetGlyGlyAsnIleThrArgValGluSerGluAsnLysValValIleLeuAspSer  
GAGATGGGCGGCAACATCACCAGGGTTGAGTCAGAAAACAAAGTGGTGATTCTGGACTCC  
CTCTACCCGCCGTTGTAGTGGTCCCACTCAGTCTTTTGTTCACCACTAAGACCTGAGG

3062 PheAspProLeuValAlaGluGluAspGluArgGluIleSerValProAlaGluIleLeu  
TTCGATCCGCTTGTGGCGGAGGAGGACGAGCGGGAGATCTCCGTACCCGCAGAAATCCTG  
AAGCTAGGCGAACACCGCCTCCTCCTGCTCGCCCTCTAGAGGCATGGGCGTCTTTAGGAC  
^

3096 BGL2,

3122 ArgLysSerArgArgPheAlaGlnAlaLeuProValTrpAlaArgProAspTyrAsnPro  
CGGAAGTCTCGGAGATTCGCCCAGGCCCTGCCCCGTTTGGGCGCGGCCGACTATAACCCC  
GCCTTCAGAGCCTCTAAGCGGGTCCGGGACGGGCAAACCCGCGCCGCGCCTGATATTGGGG  
^ ^

3143 ALWN1, 3164 EAG1 XMA3,

3182 ProLeuValGluThrTrpLysLysProAspTyrGluProProValValHisGlyCysPro  
CCGCTAGTGGAGACGTGGAAAAAGCCCCGACTACGAACCACCTGTGGTCCATGGCTGCCCC  
GGCGATCACCTCTGCACCTTTTTCGGGCTGATGCTTGGTGGACACCAGGTACCGACGGGC  
^ ^

3217 HGIE2, 3229 NCOI,

3242 LeuProProProLysSerProProValProProProArgLysLysArgThrValValLeu  
CTTCCACCTCCAAAGTCCCCTCCTGTGCCTCCGCCTCGGAAGAAGCGGACGGTGGTCCTC  
GAAGGTGGAGGTTTCAGGGGAGGACACGGAGGCGGAGCCTTCTTCGCCTGCCACCAGGAG

3302 ThrGluSerThrLeuSerThrAlaLeuAlaGluLeuAlaThrArgSerPheGlySerSer  
ACTGAATCAACCTATCTACTGCCTTGGCCGAGCTCGCCACCAGAAGCTTTGGCAGCTCC  
TGACTTAGTTGGGATAGATGACGGAACCGGCTCGAGCGGTGGTCTTCGAAACCGTCGAGG  
^ ^

3332 SACI, 3346 HIND3,

3362 SerThrSerGlyIleThrGlyAspAsnThrThrThrSerSerGluProAlaProSerGly  
TCAACTTCCGGCATTACGGGCGACAATACGACAACATCCTCTGAGCCCGCCCTTCTGGC  
AGTTGAAGGCCGTAATGCCCCGCTGTTATGCTGTTGTAGGAGACTCGGGCGGGGAAGACCG

3422 CysProProAspSerAspAlaGluSerTyrSerSerMetProProLeuGluGlyGluPro  
TGCCCCCCCCGACTCCGACGCTGAGTCCTATTCTCCATGCCCCCCTGGAGGGGGAGCCT  
ACGGGGGGGCTGAGGCTGCGACTCAGGATAAGGAGGTACGGGGGGGACCTCCCCCTCGGA  
^

3437 EAM11051,

3482 GlyAspProAspLeuSerAspGlySerTrpSerThrValSerSerGluAlaAsnAlaGlu  
GGGGATCCGGATCTTAGCGACGGGTGATGGTCAACGGTCAGTAGTGAGGCCAACCGGGAG  
CCCCTAGGCCTAGAATCGCTGCCCAGTACCAGTTGCCAGTCATCACTCCGGTTGCGCCTC  
^ ^ ^

3484 BAMHI, 3485 BSAB1, 3487 BSPE1,

3542 AspValValCysCysSerMetSerTyrSerTrpThrGlyAlaLeuValThrProCysAla  
GATGTGCTGTGCTGCTCAATGTCTTACTCTTGGACAGGCGCACTCGTCACCCCGTGCGCC  
CTACAGCACACGACGAGTTACAGAATGAGAACCTGTCCGCGTGAGCAGTGGGGCACGCGG

00221T 6472460

AlaGluGluGlnLysLeuProIleAsnAlaLeuSerAsnSerLeuLeuArgHisHisAsn  
3602 GCGGAAGAACAGAACTGCCCATCAATGCACTAAGCAACTCGTTGCTACGTCACCACAAT  
CGCCTTCTTGTCTTTGACGGGTAGTTACGTGATTCTGTTGAGCAACGATGCAGTGCGTGTTA

3662      LeuValTyrSerThrThrSerArgSerAlaCysGlnArgGlnLysLysValThrPheAsp  
TTGGTGTATTCCACCACCTCACGCAGTGCTTGCCAAAGGCAGAAGAAAGTCACATTTGAC  
AACCACATAAGGTGGTGGAGTGCGTCACGAACGGTTTCCGCTTCTCTTTCAGTGTAAACTG

3722 ArgLeuGlnValLeuAspSerHisTyrGlnAspValLeuLysGluValLysAlaAlaAla  
AGACTGCAAGTTCTGGACGCCATTACCAGGACGTACTCAAGGAGGTTAAAGCAGCGGCG  
TCTGACGTTCAAGACCTGTCGGTAATGGTCCTGCATGAGTTCCTCCAATTTTCGTCGCCGC

3782 SerLysValLysAlaAsnLeuLeuSerValGluGluAlaCysSerLeuThrProProHis  
TCAAAGTGAAGGCTAACTTGCTATCCGTAGAGGAAGCTTGCAGCCTGACGCCCCACAC  
AGTTTTCACTCCGATTGAACGATAGGCATCTCCTTCGAACGTCGGACTGCGGGGGTGTG

3842 SerAlaLysSerLysPheGlyTyrGlyAlaLysAspValArgCysHisAlaArgLysAla  
TCAGCCAAATCCAAGTTTGGTTATGGGGCAAAGACGTCCGTTGCCATGCCAGAAAGGCC  
AGTCGGTTTAGGTTCAAACCAATACCCCGTTTTCTGCAGGCAACGGTACGGTCTTTCCGG

ValThrHisIleAsnSerValTrpLysAspLeuLeuGluAspAsnValThrProIleAsp  
3902 GTAACCCACATCAACTCCGTGTGGAAAGACCTTCTGGAAGACAATGTAACACCAATAGAC  
CATTGGGTGTAGTTGAGGCACACCTTTCTGGAAGACCTTCTGTTACATTGTGGTTATCTG

3962 ThrThrIleMetAlaLysAsnGluValPheCysValGlnProGluLysGlyGlyArgLys  
ACTACCATCATGGCTAAGAACGAGGTTTTCTGCGTTCAGCCTGAGAAGGGGGGTCGTAAG  
TGATGGTAGTACCGATTCTTGCTCCAAAAGACGCAAGTCGGACTCTTCCCCCAGCATTTC

ProAlaArgLeuIleValPheProAspLeuGlyValArgValCysGluLysMetAlaLeu  
4022 CCAGCTCGTCTCATCGTGTCCCCGATCTGGGCGTGCGCGTGTGCGAAAAGATGGCTTTG  
GGTCGAGCAGAGTAGCACAAGGGGCTAGACCCGCACGCGCACACGCTTTTCTACCGAAC

4082 TyrAspValValThrLysLeuProLeuAlaValMetGlySerSerTyrGlyPheGlnTyr  
TACGACGTGGTTACAAAGCTCCCCTTGGCCGTGATGGGAAGCTCCTACGGATTCCAATAC  
ATGCTGCACCAATGTTTCGAGGGGAACCGGCACTACCCTTCGAGGATGCCTAAGGTTATG

4142 SerProGlyGlnArgValGluPheLeuValGlnAlaTrpLysSerLysLysThrProMet  
TCACCAGGACAGCGGGTTGAATTCCTCGTGCAAGCGTGGAGTCCAAGAAACCCCAATG  
AGTGGTCCTGTCGCCCAACTTAAGGAGCACGTTGCGACCTTCAGGTTCTTTTGGGGTTAC

GlyPheSerTyrAspThrArgCysPheAspSerThrValThrGluSerAspIleArgThr  
4202 GGGTTCTCGTATGATACCCGCTGCTTTGACTCCACAGTCACTGAGAGCGACATCCGTACG  
CCCAAGAGCATACTATGGGCGACGAAACTGAGGTGTCAGTGACTCTCGCTGTAGGCATGC

ArgAspGlnLeuGluGlnAlaLeuAspCysGluIleTyrGlyAlaCysTyrSerIleGlu

4862 AGGGACCAGCTTGAACAGGCCCTCGATTGCGAGATCTACGGGGCCTGCTACTCCATAGAA  
TCCCTGGTCGAACTTGTCGGGAGCTAACGCTCTAGATGCCCCGGACGATGAGGTTATCTT

4893 BGL2,

ProLeuAspLeuProProIleIleGlnArgLeuHisGlyLeuSerAlaPheSerLeuHis  
4922 CCACTGGATCTACCTCCAATCATTCAAAGACTCCATGGCCTCAGCGCATTTTCACTCCAC  
GGTGACCTAGATGGAGGTTAGTAAGTTTCTGAGGTACCGGAGTCGCGTAAAAGTGAGGTG

4954 NCOI,

4982 SerTyrSerProGlyGluIleAsnArgValAlaAlaCysLeuArgLysLeuGlyValPro  
AGTTACTCTCCAGGTGAAATCAATAGGGTGGCCGCATGCCTCAGAAAACCTGGGGTACCG  
TCAATGAGAGGTCCACTTTAGTTATCCACCGGCGTACGGAGTCTTTTGAACCCCATGGC

5015 SPHI, 5035 KPNI.

ProLeuArgAlaTrpArgHisArgAlaArgSerValArgAlaArgLeuLeuAlaArgGly  
5042 CCCTTGCGAGCTTGAGACACCGGGCCGAGCGTCCGCGCTAGGCTTCTGGCCAGAGGA  
GGGAACGCTCGAACCTCTGTGGCCCGGGCCTCGCAGGCGCGATCCGAAGACCGGTCTCCT

5064 APAI, 5091 BALI,

5102 GlyArgAlaAlaIleCysGlyLysTyrLeuPheAsnTrpAlaValArgThrLysLeuLys  
GGCAGGGCTGCCATATGTGGCAAGTACCTCTTCAACTGGGCAGTAAGAACAAAGCTCAAA  
CCGTCCCGACGGTATACACCGTTCATGGAGAAGTTGACCCGTCATTCTTGTTCGAGTTT

5113 NDEI,

LeuThrProIleAlaAlaAlaGlyGlnLeuAspLeuSerGlyTrpPheThrAlaGlyTyr  
5162 CTCACTCCAATAGCGGCCGCTGGCCAGCTGGACTTGTCCGGTGTTTACGGCTGGCTAC  
GAGTGAGGTTATCGCCGGCACCAGGTCGACCTGAACAGGCCGACCAAGTGCCGACCGATG

5174 NOTI, 5175 EAG1 XMA3, 5182 BALI, 5186 PVU2.

5222 SerGlyGlyAspIleTyrHisSerValSerHisAlaArgProArgTrpIleTrpPheCys  
AGCGGGGGAGACATTTATCACAGCGTGTCTCATGCCCGGCCCGCTGGATCTGGTTTTGC  
TCGCCCCCTCTGTAAATAGTGTGCGACAGAGTACGGGGCGGGGCGACCTAGACCAAACG

5240 DRA3,

LeuLeuLeuLeuAlaAlaGlyValGlyIleTyrLeuLeuProAsnArgMetSerThrAsn  
5282 CTACTCCTGCTTGCTGCAGGGGTAGGCATCTACCTCCTCCCCAACCGAATGAGCACGAAT  
GATGAGGACGAACGACGTCCCCATCCGTAGATGGAGGAGGGGTTGGCTTACTCGTGCTTA

5295 PSTI,

ProLysProGlnArgLysThrLysArgAsnThrAsnArgArgProGlnAspValLysPhe  
5342 CCTAAACCTCAAAGAAAGACCAACGTAACACCAACCGCGCGCCGAGGACGTCAAGTTC  
GGATTTGGAGTTTCTTTCTGGTTTGCATTGTGGTTGGCCGCCGGCGTCTCTGCAGTTCAAG

5380 NOT1, 5381 EAG1 XMA3, 5390 AAT2, 5401 SMAI XMAI,

ProGlyGlyGlyGlnIleValGlyGlyValTyrLeuLeuProArgArgGlyProArgLeu  
5402 CCGGGTGGCGGT CAGATCGTTGGTGGAGTTTACTTGTTCGCCGCAGGGGCCCTAGATTG  
GGCCCCACGCCAGTCTAGCAACCACTCAAATGAACAACGGCGCGTCCCCGGGATCTAAC



5449 APAI,

5462 GlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGlyArgArgGlnPro  
GGTGTGCGCGCGACGAGAAAGACTTCCGAGCGGTGCGAACCTCGAGGTAGACGTCAGCCT  
CCACACGCGCGCTGCTCTTTCTGAAGGCTCGCCAGCGTTGGAGCTCCATCTGCAGTCGGA

5467 BSSH2, 5478 XMNI, 5502 XHOI, 5511 AAT2,

5522 IleProLysAlaArgArgProGluGlyArgThrTrpAlaGlnProGlyTyrProTrpPro  
ATCCCCAAGGCTCGTCGGCCCGAGGGCAGGACCTGGGCTCAGCCCGGGTACCCTTGGCCC  
TAGGGGTTCCGAGCAGCCGGGCTCCCGTCCTGGACCCGAGTCGGGGCCCATGGGAACCGGG

5548 ALWN1, 5558 ESP1, 5564 SMAI XMAI, 5568 KPNI,

5582 LeuTyrGlyAsnGluGlyCysGlyTrpAlaGlyTrpLeuLeuSerProArgGlySerArg  
CTCTATGGCAATGAGGGCTGCGGGTGGGCGGGATGGCTCCTGTCTCCCGTGGCTCTCGG  
GAGATACCGTTACTCCCGACGCCCCACCCGCCCTACCGAGGACAGAGGGGCACCGAGAGCC

5642 ProSerTrpGlyProThrAspProArgArgArgSerArgAsnLeuGlyLysValIleAsp  
CCTAGCTGGGGCCCCACAGACCCCGGCGTAGGTGCGCAATTTGGGTAAGGTCATCGAT  
GGATCGACCCCGGGGTGTCTGGGGGCCGCATCCAGCGCGTTAAACCCATTCCAGTAGCTA

5650 APAI, 5696 CLAI,

5702 ThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIleProLeuValOC AM  
ACCCTTACGTGCGGCTTCGCGGACCTCATGGGGTACATACCGCTCGTCTAATAGTCGAC  
TGGGAATGCACGCCGAAGCGGCTGGAGTACCCCATGTATGGCGAGCAGATTATCAGCTG

5724 HGIE2, 5755 SALI,

644T260

MetAlaAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsn  
2 AGCTTACAAAACAAAATGGCTGCATATGCAGCTCAGGGCTATAAGGTGCTAGTACTCAAC  
TCGAATGTTTTGTTTTACCGACGTATACGTCGAGTCCCGATATTCCACGATCATGAGTTG  
1 HIND3, 24 NDEI, 52 SCAI,  
ProSerValAlaAlaThrLeuGlyPheGlyAlaTyrMetSerLysAlaHisGlyIleAsp  
62 CCCTCTGTTGCTGCAACACTGGGCTTTGGTGCTTACATGTCCAAGGCTCATGGGATCGAT  
GGGAGACAACGACGTTGTGACCCGAAACCACGAATGTACAGGTTCCGAGTACCCTAGCTA  
116 CLAI,  
ProAsnIleArgThrGlyValArgThrIleThrThrGlySerProIleThrTyrSerThr  
122 CCTAACATCAGGACCGGGGTGAGAACAATTACCACTGGCAGCCCCATCACGTACTCCACC  
GGATTGTAGTCCTGGCCCCACTCTTGTTAATGGTGACCGTCGGGGTAGTGCATGAGGTGG  
TyrGlyLysPheLeuAlaAspGlyGlyCysSerGlyGlyAlaTyrAspIleIleIleCys  
182 TACGGCAAGTTCCTTGCCGACGGCGGGTGCTCGGGGGGCGCTTATGACATAATAATTTGT  
ATGCCGTTCAAGGAACGGCTGCCGCCACGAGCCCCCGCGAATACTGTATTATTAAACA  
AspGluCysHisSerThrAspAlaThrSerIleLeuGlyIleGlyThrValLeuAspGln  
242 GACGAGTGCCACTCCACGGATGCCACATCCATCTTGGGCATTGGCACTGTCCTTGACCAA  
CTGCTCACGGTGAGGTGCCTACGGTGTAGGTAGAACCCGTAACCGTGACAGGAAGTGGT  
AlaGluThrAlaGlyAlaArgLeuValValLeuAlaThrAlaThrProProGlySerVal  
302 GCAGAGACTGCGGGGCGGAGACTGGTTGTGCTCGCCACCGCCACCCCTCCGGGCTCCGTC  
CGTCTCTGACGCCCCCGCTCTGACCAACACGAGCGGTGGCGGTGGGGAGGCCCGAGGCAG  
303 ALWN1,  
ThrValProHisProAsnIleGluGluValAlaLeuSerThrThrGlyGluIleProPhe  
362 ACTGTGCCCATCCCAACATCGAGGAGGTTGCTCTGTCCACCACCGAGAGATCCCTTTT  
TGACACGGGGTAGGGTTGTAGCTCCTCCAACGAGACAGGTGGTGGCCTCTCTAGGGAAAA  
TyrGlyLysAlaIleProLeuGluValIleLysGlyGlyArgHisLeuIlePheCysHis  
422 TACGGCAAGGCTATCCCCCTCGAAGTAATCAAGGGGGGAGACATCTCATCTTCTGTGTCAT  
ATGCCGTTCCGATAGGGGGAGCTTCATTAGTTCCCCCCTCTGTAGAGTAGAAGACAGTA

SerLysLysLysCysAspGluLeuAlaAlaLysLeuValAlaLeuGlyIleAsnAlaVal  
 482 TCAAAGAAGAAGTGGCGACGAAGCTCGCCGCAAAGCTGGTCGCATTGGGCATCAATGCCGTG  
 AGTTTCTTCTTACGCTGCTTGAGCGGCGTTTCGACCAGCGTAACCCGTAGTTACGGCAC  
  
 AlaTyrTyrArgGlyLeuAspValSerValIleProThrSerGlyAspValValValVal  
 542 GCCTACTACCGCGGTCTTGACGTGTCCGTCATCCCGACCAGCGGCGATGTTGTGTCGTG  
 CGGATGATGGCGCCAGAACTGCACAGGCAGTAGGGCTGGTCGCCGCTACAACAGCAGCAC  
 ^ ^  
 550 SAC2, 560 DRD1,  
  
 AlaThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsn  
 602 GCAACCGATGCCCTCATGACCGGCTATACCGGCGACTTCGACTCGGTGATAGACTGCAAT  
 CGTTGGCTACGGGAGTACTGGCCGATATGGCCGCTGAAGCTGAGCCACTATCTGACGTTA  
 ^  
 615 BSPH1,  
  
 ThrCysValThrGlnThrValAspPheSerLeuAspProThrPheThrIleGluThrIle  
 662 ACGTGTGTCAACCCAGACAGTCGATTTTCAGCCTTGACCCTACCTTCACCATTGAGACAATC  
 TGCACACAGTGGGTCTGTCTAGCTAAAGTCGGAAGTGGGATGGAAGTGGTAACCTCTGTTAG  
  
 ThrLeuProGlnAspAlaValSerArgThrGlnArgArgGlyArgThrGlyArgGlyLys  
 722 ACGTCCCCCAAGATGCTGTCTCCCGCACTCAACGTCGGGGCAGGACTGGCAGGGGGAAG  
 TCGAGGGGGTTCTACGACAGAGGGCGTGAGTTGCAGCCCCGTCCTGACCGTCCCCCTTC  
  
 ProGlyIleTyrArgPheValAlaProGlyGluArgProSerGlyMetPheAspSerSer  
 782 CCAGGCATCTACAGATTTGTGGCACCGGGGGAGCGCCCCCTCCGGCATGTTGACTCGTCC  
 GGTCCGTAGATGTCTAAACACCGTGGCCCCCTCGCGGGGAGGCCGTACAAGCTGAGCAGG  
 ^ ^  
 816 BGLI, 833 DRD1,  
  
 ValLeuCysGluCysTyrAspAlaGlyCysAlaTrpTyrGluLeuThrProAlaGluThr  
 842 GTCCTCTGTGAGTGCTATGACGCAGGCTGTGCTTGGTATGAGCTCACGCCCCGCGAGACT  
 CAGGAGACACTCACGATACTGCGTCCGACACGAACCATACTCGAGTGCGGGCGGCTCTGA  
 ^  
 881 SACI,  
  
 ThrValArgLeuArgAlaTyrMetAsnThrProGlyLeuProValCysGlnAspHisLeu  
 902 ACAGTTAGGCTACGAGCGTACATGAACACCCCGGGGCTTCCCGTGTGCCAGGACCATCTT  
 TGTCATCCGATGCTCGCATGTACTTGTGGGGCCCCGAAGGGCACACGGTCCTGGTAGAA  
 ^  
 931 SMAI XMAI,  
  
 GluPheTrpGluGlyValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGln  
 962 GAATTTTGGGAGGGCGTCTTTACAGGCCTCACTCATATAGATGCCCACTTTCTATCCCAG  
 CTTAAACCCTCCCGCAGAAATGTCCGGAGTGAGTATATCTACGGGTGAAAGATAGGGTC  
 ^  
 985 STUI,  
  
 ThrLysGlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGlnAlaThrValCysAla  
 1022 ACAAAGCAGAGTGGGGAGAACCTTCCTTACCTGGTAGCGTACCAAGCCACCGTGTGCGCT  
 TGTTTTCGTCTACCCCTCTTGGAAGGAATGGACCATCGCATGGTTCGGTGGCACACGCGA  
 ^  
 1069 DRA3,  
  
 ArgAlaGlnAlaProProProSerTrpAspGlnMetTrpLysCysLeuIleArgLeuLys  
 1082 AGGGCTCAAGCCCCTCCCCCATCGTGGGACCAGATGTGGAAGTGTGTTGATTGCCTCAAG

002217 6472450

AlaValThrSerProLeuThrThrSerGlnThrLeuLeuPheAsnIleLeuGlyGlyTrp  
1682 GCTGTCACCAGCCCACTAACCCTAGCCAAACCCTCCTCTTCAACATATTGGGGGGGTGG  
CGACAGTGGTCGGGTGATTGGTGATCGGTTTGGGAGGAGAAGTTGTATAACCCCCCACC

# FIGURE 22 - Page 4

ValAlaAlaGlnLeuAlaAlaProGlyAlaAlaThrAlaPheValGlyAlaGlyLeuAla  
 1742 GTGGCTGCCCAGCTCGCCGCCCCCGGTGCCGCTACTGCCTTTGTGGGCGCTGGCTTAGCT  
 CACCGACGGGTCGAGCGGGGGGCCACGGCGATGACGGAAACACCCGCGACCGAATCGA  
 ^

1794 ESP1,

GlyAlaAlaIleGlySerValGlyLeuGlyLysValLeuIleAspIleLeuAlaGlyTyr  
 1802 GGCGCCGCCATCGGCAGTGTGGACTGGGGAAGGTCCTCATAGACATCCTTGCAGGGTAT  
 CCGCGGCGGTAGCCGTACAACTGACCCCTTCCAGGAGTATCTGTAGGAACGTCCCATA  
 ^

1802 KAS1 NARI,

GlyAlaGlyValAlaGlyAlaLeuValAlaPheLysIleMetSerGlyGluValProSer  
 1862 GGCGCGGGCGTGGCGGGAGCTCTTGTGGCATTCAAGATCATGAGCGGTGAGGTCCCTCC  
 CCGCGCCCGCACCGCCCTCGAGAACACCGTAAGTTCTAGTACTCGCCACTCCAGGGGAGG  
 ^

1878 SACI, 1899 BSPH1,

ThrGluAspLeuValAsnLeuLeuProAlaIleLeuSerProGlyAlaLeuValValGly  
 1922 ACGGAGGACCTGGTCAATCTACTGCCCCGCCATCCTCTCGCCCGAGCCCTCGTAGTCGGC  
 TGCCTCTCTGGACCAGTTAGATGACGGGCGGTAGGAGAGCGGGCCTCGGGAGCATCAGCCG  
 ^

1928 TTH3I,

ValValCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaValGlnTrp  
 1982 GTGGTCTGTGCAGCAATACTGCGCCGGCACGTTGGCCCGGGCGAGGGGCGAGTGCAGTGG  
 CACCAGACACGTCGTTATGACGCGGCCGTGCAACCGGGCCCGTCCCCCGTCACGTCACC  
 ^

2004 NAEI, 2017 SMAI XMAI,

MetAsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValSerProThrHisTyrVal  
 2042 ATGAACCGGCTGATAGCCTTCGCCTCCCGGGGAACCATGTTTCCCCACGCACTACGTG  
 TACTTGGCCGACTATCGGAAGCGGAGGGCCCCCTTGGTACAAAGGGGGTGCGTGATGCAC  
 ^

2067 SMAI XMAI, 2093 DRA3,

ProGluSerAspAlaAlaAlaArgValThrAlaIleLeuSerSerLeuThrValThrGln  
 2102 CCGGAGAGCGATGCAGCTGCCCGCGTCACTGCCATACTCAGCAGCCTCACTGTAACCCAG  
 GGCTCTCGCTACGTGACGGGCGCAGTGACGGTATGAGTCGTCGGAGTGACATTGGGTC  
 ^

2115 PVU2, 2159 ALWN1,

LeuLeuArgArgLeuHisGlnTrpIleSerSerGluCysThrThrProCysSerGlySer  
 2162 CTCCTGAGGCGACTGCACCACTGGATAAGCTCGGAGTGTACCACTCCATGCTCCGGTTCC  
 GAGGACTCCGCTGACGTGGTCACCTATTCGAGCCTCACATGGTGAGGTACGAGGCCAAGG  
 ^

2164 MST2, 2220 ECON1,

TrpLeuArgAspIleTrpAspTrpIleCysGluValLeuSerAspPheLysThrTrpLeu  
 2222 TGGCTAAGGGACATCTGGGACTGGATATGCGAGGTGTTGAGCGACTTTAAGACCTGGCTA  
 ACCGATTCCCTGTAGACCCTGACCTATACGCTCCACAACCTCGCTGAAATTCTGGACCGAT

LysAlaLysLeuMetProGlnLeuProGlyIleProPheValSerCysGlnArgGlyTyr  
 2282 AAAGCTAAGCTCATGCCACAGCTGCCTGGGATCCCCCTTTGTGTCCTGCCAGCGCGGGTAT  
 TTTCGATTGAGTACGGTGTGACGGACCCTAGGGGAAACACAGGACGGTTCGCGCCCAT  
 ^

2285 ESP1, 2300 PVU2, 2310 BAMHI,

| Table 1. Summary of the data used in the study. |                                                                                                                                                                                                                                       |
|-------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Study                                           | 1997-2000                                                                                                                                                                                                                             |
| Location                                        | United States                                                                                                                                                                                                                         |
| Sample size                                     | 1,000,000                                                                                                                                                                                                                             |
| Response rate                                   | 85%                                                                                                                                                                                                                                   |
| Age range                                       | 18-80 years                                                                                                                                                                                                                           |
| Gender                                          | Male: 50%, Female: 50%                                                                                                                                                                                                                |
| Ethnicity                                       | White: 60%, Black: 20%, Hispanic: 15%, Asian: 5%                                                                                                                                                                                      |
| Education                                       | High school or less: 30%, College: 40%, Graduate: 30%                                                                                                                                                                                 |
| Income                                          | \$0-\$10,000: 20%, \$10,000-\$20,000: 25%, \$20,000-\$30,000: 20%, \$30,000-\$40,000: 15%, \$40,000-\$50,000: 10%, \$50,000-\$60,000: 5%, \$60,000-\$70,000: 3%, \$70,000-\$80,000: 2%, \$80,000-\$90,000: 1%, \$90,000-\$100,000: 1% |
| Health status                                   | Good: 60%, Fair: 20%, Poor: 20%                                                                                                                                                                                                       |
| Insurance                                       | Medicare: 15%, Medicaid: 10%, Private: 75%                                                                                                                                                                                            |
| Employment                                      | Employed: 60%, Unemployed: 40%                                                                                                                                                                                                        |
| Marital status                                  | Married: 50%, Single: 30%, Divorced: 15%, Widowed: 5%                                                                                                                                                                                 |
| Living arrangement                              | Alone: 30%, With family: 40%, With friends: 10%, In nursing home: 20%                                                                                                                                                                 |
| Comorbidities                                   | Hypertension: 25%, Diabetes: 15%, Heart disease: 10%, Stroke: 5%, Cancer: 10%, Asthma: 5%, Depression: 10%, Anxiety: 5%                                                                                                               |
| Medication use                                  | Antidepressants: 10%, Antipsychotics: 5%, Anticonvulsants: 5%, Beta-blockers: 10%, Calcium channel blockers: 10%, Diuretics: 10%, Insulin: 5%, Lipid-lowering agents: 10%, Painkillers: 10%, Sedatives: 5%                            |
| Healthcare utilization                          | Physician visits: 10%, Hospitalizations: 5%, Emergency department visits: 5%, Long-term care: 5%                                                                                                                                      |
| Quality of life                                 | Physical: 50%, Mental: 40%, Social: 30%                                                                                                                                                                                               |
| Healthcare costs                                | \$0-\$10,000: 20%, \$10,000-\$20,000: 25%, \$20,000-\$30,000: 20%, \$30,000-\$40,000: 15%, \$40,000-\$50,000: 10%, \$50,000-\$60,000: 5%, \$60,000-\$70,000: 3%, \$70,000-\$80,000: 2%, \$80,000-\$90,000: 1%, \$90,000-\$100,000: 1% |

| Table 1. Demographic characteristics of the study population |                |
|--------------------------------------------------------------|----------------|
| Age (years)                                                  | 65.0 ± 10.0    |
| Gender                                                       |                |
| Male                                                         | 50 (50.0%)     |
| Female                                                       | 50 (50.0%)     |
| Education (years)                                            | 12.0 ± 2.0     |
| Marital status                                               |                |
| Married                                                      | 40 (80.0%)     |
| Single                                                       | 10 (20.0%)     |
| Occupation                                                   |                |
| Retired                                                      | 30 (60.0%)     |
| Unemployed                                                   | 20 (40.0%)     |
| Income (USD/month)                                           | 1000.0 ± 200.0 |
| Health status                                                |                |
| Good                                                         | 40 (80.0%)     |
| Poor                                                         | 10 (20.0%)     |
| Comorbidities                                                |                |
| Hypertension                                                 | 20 (40.0%)     |
| Diabetes                                                     | 10 (20.0%)     |
| Cholesterol                                                  | 15 (30.0%)     |
| Smoking status                                               |                |
| Smoker                                                       | 10 (20.0%)     |
| Non-smoker                                                   | 40 (80.0%)     |
| Alcohol consumption                                          |                |
| Regular                                                      | 5 (10.0%)      |
| Occasional                                                   | 15 (30.0%)     |
| Never                                                        | 30 (60.0%)     |

3589 DRA3, 3600 SAC2, -

AlaGluGluGlnLysLeuProIleAsnAlaLeuSerAsnSerLeuLeuArgHisHisAsn  
3602 GCGGAAGAACAGAACTGCCCATCAATGCACTAAGCAACTCGTTGCTACGTCACCACAAT  
CGCCTTCTTGCTTTGACGGGTAGTTACGTGATTGCTTGAGCAACGATGCAGTGGTGTTA

3611 ALWN1, 3655 PFLM1,

3662    LeuValTyrSerThrThrSerArgSerAlaCysGlnArgGlnLysLysValThrPheAsp  
          TTGGTGTATTCCACCACCTCACGCAGTGCTTGCCAAAGGCAGAAGAAAGTCACATTTGAC  
          AACCACATAAGGTGGTGGAGTGCCTCACGAACGGTTTTCCGTCTTCTTTTCAGTGTAAACTG

3681 DRA3,

3722 ArgLeuGlnValLeuAspSerHisTyrGlnAspValLeuLysGluValLysAlaAlaAla  
AGACTGCAAGTTCTGGACAGCCATTACCAGGACGTACTCAAGGAGGTTAAAGCAGCGGCG  
TCTGACGTTCAAGACCTGTCGGTAATGGTCCTGCATGAGTTCTCCAATTTTCGTCGCCGC

3782 SerLysValLysAlaAsnLeuLeuSerValGluGluAlaCysSerLeuThrProProHis  
TCAAAGTGAAAGGCTAACTTGCTATCCGTAGAGGAAGCTTGACAGCTGACGCCCCACAC  
AGTTTTCACTTCCGATTGAACGATAGGCATCTCCTTCGAACGTCGGACTGCGGGGGTGTG

3816 HIND3,

3842 SerAlaLysSerLysPheGlyTyrGlyAlaLysAspValArgCysHisAlaArgLysAla  
TCAGCCAAATCCAAGTTTGGTTATGGGGCAAAGACGTCCGTTGCCATGCCAGAAAGGCC  
AGTCGGTTTAGGTTCAAACCAATACCCCGTTTTCTGCAGGCAACGGTACGGTCTTTCCGG

3875 AAT2, 3890 BGLI,

ValThrHisIleAsnSerValTrpLysAspLeuLeuGluAspAsnValThrProIleAsp  
3902 GTAACCCACATCAACTCCGTGTGGAAAGACCTTCTGGAAGACAATGTAACACCAATAGAC  
CATTGGGTGTAGTTGAGGCACACCTTTCTGGAAGACCTTCTGTTACATTGTGGTTATCTG

3962 ThrThrIleMetAlaLysAsnGluValPheCysValGlnProGluLysGlyGlyArgLys  
ACTACCATCATGGCTAAGAACGAGGTTTTCTGCGTTCAGCCTGAGAAGGGGGGTCGTAAG  
TGATGGTAGTACCGATTCTTGCTCCAAAGACGCAAGTCGGACTCTTCCCCCAGCATTC

ProAlaArgLeuIleValPheProAspLeuGlyValArgValCysGluLysMetAlaLeu  
4022 CCAGCTCGTCTCATCGTGTCCCCGATCTGGGCGTGCGGTGTGCGAAAAGATGGCTTTG  
GGTCGAGCAGAGTAGCACAAGGGGCTAGACCCGCACGCGCACACGCTTTTCTACCGAAAC

4082 TyrAspValValThrLysLeuProLeuAlaValMetGlySerSerTyrGlyPheGlnTyr  
TACGACGTGGTTACAAAGCTCCCCTTGCCGTGATGGGAAGCTCCTACGGATTCCAATAC  
ATGCTGCACCAATGTTTCGAGGGGAACCGGCACTACCCCTTCGAGGATGCCTAAGGTTATG

4142 SerProGlyGlnArgValGluPheLeuValGlnAlaTrpLysSerLysLysThrProMet  
TCACCAGGACAGCGGGTTGAATTCCTCGTGCAAGCGTGGAGTCCAAGAAACCCCAATG  
AGTGGTCCTGTCGCCCACTTAAGGAGCACGTTGCGACCTTCAGGTTCTTTTGGGGTTAC

4160 ECORI,

4202 GlyPheSerTyrAspThrArgCysPheAspSerThrValThrGluSerAspIleArgThr  
GGGTTCTCGTATGATACCCGCTGCTTTGACTCCACAGTCACTGAGAGCGACATCCGTACG  
CCCAAGAGCATACTATGGGCGACGAACTGAGGTGTCAGTGACTCTCGCTGTAGGCATGC





4862 AGGGACCAGCTTGAACAGGCCCTCGATTGCGAGATCTACGGGGCCTGCTACTCCATAGAA  
TCCCTGGTCGAACTTGTCCGGGAGCTAACGCTCTAGATGCCCCGGACGATGAGGTATCTT

4893 BGL2,

ProLeuAspLeuProProIleIleGlnArgLeuHisGlyLeuSerAlaPheSerLeuHis  
4922 CCACTGGATCTACCTCCAATCATTCAAAGACTCCATGGCCTCAGCGCATTTTCACTCCAC  
GGTGACCTAGATGGAGGTTAGTAAGTTTCTGAGGTACCGGAGTCGCGTAAAAGTGAGGTG

4954 NCOI,

SerTyrSerProGlyGluIleAsnArgValAlaAlaCysLeuArgLysLeuGlyValPro  
4982 AGTTACTCTCCAGGTGAAATCAATAGGTTGGCCGCATGCCTCAGAAAACCTTGGGGTACCG  
TCAATGAGAGGTCCACTTTAGTTATCCACCGCGGTACGGAGTCTTTTGAACCCCATGGC

5015 SPHI, 5035 KPNI,

ProLeuArgAlaTrpArgHisArgAlaArgSerValArgAlaArgLeuLeuAlaArgGly  
5042 CCCTTGCGAGCTTGAGACACCGGGCCCGGAGCGTCCGCGCTAGGCTTCTGGCCAGAGGA  
GGGAACGCTCGAACCTCTGTGGCCCGGCCTCGCAGGCGCGATCCGAAGACCGGTCTCTCT

5064 APAI, 5091 BALI,

GlyArgAlaAlaIleCysGlyLysTyrLeuPheAsnTrpAlaValArgThrLysLeuLys  
5102 GGCAGGGCTGCCATATGTGGCAAGTACCTCTTCAACTGGGCAGTAAGAACAAAGCTCAAA  
CCGTCCCGACGGTATACACCGTTCATGGAGAAGTTGACCCGTCATTCTTGTTCGAGTTT

5113 NDEI,

LeuThrProIleAlaAlaAlaGlyGlnLeuAspLeuSerGlyTrpPheThrAlaGlyTyr  
5162 CTCACTCCAATAGCGGCCGCTGGCCAGCTGGACTTGTCCGGCTGGTTTCACGGCTGGCTAC  
GAGTGAGGTTATCGCCGCGACCGGTGACCTGAACAGGCCGACCAAGTGCCGACCGATG

5174 NOTI, 5175 EAGI XMA3, 5182 BALI, 5186 PVU2,

SerGlyGlyAspIleTyrHisSerValSerHisAlaArgProArgTrpIleTrpPheCys  
5222 AGCGGGGGAGACATTTATCACAGCGTGTCTCATGCCCGGCCCGCTGGATCTGGTTTTGC  
TCGCCCCCTCTGTAAATAGTGTGCGACAGAGTACGGGCCGGGGCGACCTAGACCAAACG

5240 DRA3,

LeuLeuLeuLeuAlaAlaGlyValGlyIleTyrLeuLeuProAsnArgMetSerThrAsn  
5282 CTACTCCTGCTTGCTGCAGGGGTAGGCATCTACCTCCTCCCAACCGAATGAGCACGAAT  
GATGAGGACGAACGACGTCCCCATCCGTAGATGGAGGAGGGGTTGGCTTACTCGTGCTTA

5295 PSTI,

ProLysProGlnArgLysThrLysArgAsnThrAsnArgArgProGlnAspValLysPhe  
5342 CCTAAACCTCAAAGAAAGACCAAACGTAACACCAACCGCGGCCGAGGACGTCAAGTTC  
GGATTTGGAGTTTCTTTCTGGTTTGCATTGTGGTTGGCCGCCGCGTCCTGCAGTTCAAG

5380 NOTI, 5381 EAGI XMA3, 5390 AAT2, 5401 SMAI XMAI,

ProGlyGlyGlyGlnIleValGlyGlyValTyrLeuLeuProArgArgGlyProArgLeu  
5402 CCGGGTGGCGGTCAGATCGTTGGTGGAGTTTACTTGTGCGCGCAGGGGCCCTAGATTG  
GGCCACCGCCAGTCTAGCAACCACCTCAAATGAACAACGGCGCGTCCCGGGATCTAAC

5449 APAI,

5462 GlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGlyArgArgGlnPro  
GGTGTGCGCGCGACGAGAAAGACTTCCGAGCGGTGCGAACCTCGAGGTAGACGTACGCCT  
CCACACGCGCGCTGCTCTTTCTGAAGGCTCGCCAGCGTTGGAGCTCCATCTGCAGTCGGA

5467 BSSH2, 5478 XMNI, 5502 XHOI, 5511 AAT2,

5522 IleProLysAlaArgArgProGluGlyArgThrTrpAlaGlnProGlyTyrProTrpPro  
ATCCCCAAGGCTCGTCGGCCCCAGGGCAGGACCTGGGCTCAGCCCGGTACCTTGGCCC  
TAGGGGTTCGAGCAGCCGGGCTCCCGTCCTGGACCCAGTCGGGCCCATGGGAACCGGG

5548 ALWN1, 5558 ESP1, 5564 SMAI XMAI, 5568 KPNI,

LeuTyrGlyAsnGluGlyCysGlyTrpAlaGlyTrpLeuLeuSerProArgGlySerArg  
5582 CTCTATGGCAATGAGGGCTGCGGGTGGGCGGGATGGCTCTGTCTCCCGTGGCTCTCGG  
GAGATACCGTTACTCCCGACGCCCCACCCGCCCTACCGAGGACAGAGGGGCACCGAGAGCC

ProSerTrpGlyProThrAspProArgArgArgSerArgAsnLeuGlyLysValIleAsp  
5642 CCTAGCTGGGGCCCCACAGACCCCGGCGTAGGTCGCGCAATTTGGGTAAGGTCATCGAT  
GGATCGACCCCGGGGTGTCTGGGGGCCGCATCCAGCGCGTTAAACCCATTCCAGTAGCTA

5650 APAI, 5696 CLAI,

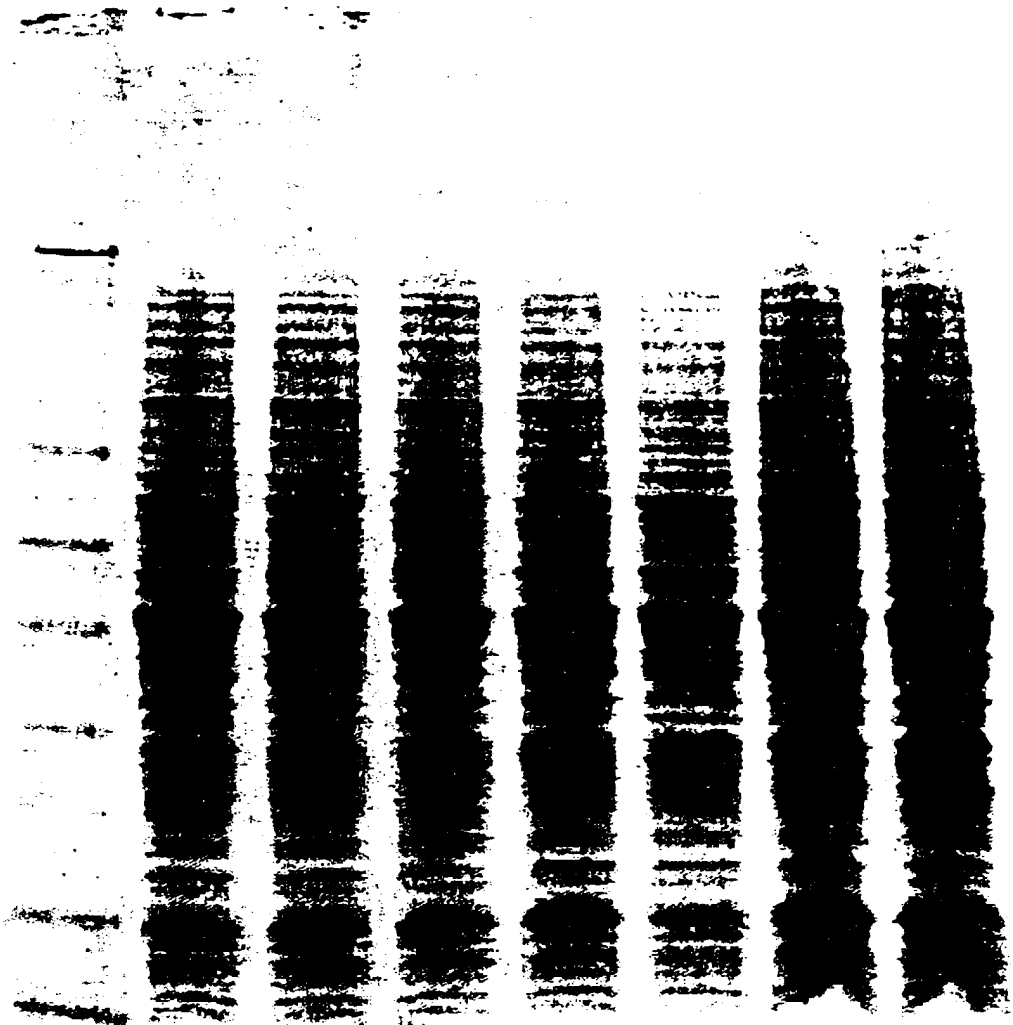
5702 ThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIleProLeuValGlyAlaProLeu  
ACCCTTACGTGCGGCTTCGCCGACCTCATGGGGTACATACCGCTCGTCGGCGCCCTCTT  
TGGGAATGCACGCCGAAGCGGCTGGAGTACCCCATGTATGGCAGCAGCCGCGGGGAGAA

5724 HGIE2, 5750 KAS1 NARI, 5756 ECON1,

GlyGlyAlaAlaArgAlaOC AM  
5762 GGAGGCGCTGCCAGGGCCTAATAGTCGAC  
CCTCCGCGACGGTCCCGGATTATCAGCTG

5785 SALI,

FIGURE 23



002275 5247260